March 2004

The Starry Night Texture

Xenophon Zabulis
University of Pennsylvania

Benjamin T. Backus
University of Pennsylvania, backus@psych.upenn.edu

University of Pennsylvania Institute for Research in Cognitive Science Technical Report No. IRCS-04-01.

The Starry Night Texture

Abstract

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 stimulus based on a novel texture, the Starry Night Texture (SNT), that is incapable of conveying slant, depth edges, or texture boundaries, in a single view. Changing density and changing intensity are equivalent for SNT, so an instance of the texture is characterized (up to the random locations of the texture elements) by its densintensity. We describe the SNT in its ideal form, consider deviations from the ideal that are needed to realize the texture in practice, and describe a physical device that approximates SNT using backlit metal foil. In three experiments with computer-generated stimuli we examined human perception of SNT, to show that (1) the deviations from 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 better than other textures across changes in distance; and (3) the use of SNT in a slant perception experiment did not significantly increase observers' reliance on stereoscopic slant cues, as compared to the sparse random dot displays that have been commonly employed to study human perception of shape from binocular disparity and motion.

Keywords

Texture, stereoscopic vision, cue combination

Comments

University of Pennsylvania Institute for Research in Cognitive Science Technical Report No. IRCS-04-01.

The Starry Night Texture

Xenophon Zabulis[1,2] and Benjamin T. Backus[1,3]

[1]Institute for Research in Cognitive Science
[2]GRASP Laboratory
[3]Department of Psychology

University of Pennsylvania

IRCS Technical Report
11 March 2004

Keywords: Texture, stereoscopic vision, cue combination

Correspondence: B. T. Backus
3401 Walnut St., Room 302A, Philadelphia, PA 19104-6228
backus@psych.upenn.edu
phone 215-573-9341

ABSTRACT

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 stimulus based on a novel texture, the Starry Night Texture (SNT), that is incapable of conveying slant, depth edges, or texture boundaries, in a single view. Changing density and changing intensity are equivalent for SNT, so an instance of the texture is characterized (up to the random locations of the texture elements) by its *densintensity*. We describe the SNT in its ideal form, consider deviations from the ideal that are needed to realize the texture in practice, and describe a physical device that approximates SNT using backlit metal foil. In three experiments with computer-generated stimuli we examined human perception of SNT, to show that (1) the deviations from 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 better than other textures across changes in distance; and (3) the use of SNT in a slant perception experiment did not significantly increase observers' reliance on stereoscopic slant cues, as compared to the sparse random dot displays that have been commonly employed to study human perception of shape from binocular disparity and motion.

1. Introduction

Consider a Julesz random-dot stereogram (RDS),[1] 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] If head movements are allowed, then the absence of depth edges is also specified by motion parallax cues. Thus, Julesz's observation, that binocular disparity per se is sufficient to evoke depth, is actually more impressive when one realizes the nature of this conflict. 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. We will not be concerned here with contrived, inhomogeneous textures that mislead the observer as to depth structure. That is because texture cues are probabilistic by nature: any depth structure could give rise to any image, if the structure were painted with the

proper inhomogeneous surface pattern.

The top row of Figure 3 illustrates two examples of depth information conveyed by texture. A homogeneously patterned surface is shown at the top left. When the distance to this surface is reduced by a factor of two (top middle panel), dot sizes and spacing increase, from which an observer could infer that the distance has been halved. No matching of individual texture elements is necessary to make this inference; it is the change in the image statistics (dot size and density) that specify the change in distance. Similarly, if the surface is slanted by 60° (top right panel), its image will contain a giveaway texture gradient and, in this case, linear perspective and outline cues. If one assumes homogeneity, then changes in slant across an imaged surface can be quantified by fitting affine transformations from one part of the image to another;[4,5] a similar scheme could reveal changes in distance (up to a scale factor) across depth edges. Even similarly textured surfaces that abut, without any change in depth across their edge, usually give rise to a detectable edge between them, that appears as a "second order" contour within the image. The orderly layout of texture elements may suffer a discontinuity, or extended texture elements may themselves be cut by the edge (not shown).

Thus, a texture must be neither homogeneous, nor contain identifiable texture elements, if it is to completely hide slant, depth edges, and edges between abutting surfaces. These considerations dictate that the texture must be comprised of 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 the case of small luminous dots, changes in depth also give rise to variation in the "proximity luminance" cue.[6] To completely hide changes in depth, one would have to make it impossible to reliably compute not only the transformations that relate patches in different parts of the image to one another, but also impossible to measure changes in dot density and intensity.

The Starry Night Texture is a random-dot texture that, in its ideal form, renders depth edges and changes in surface slant invisible within the image. Four properties of SNT confer it with this ability; the third and fourth of these are novel to SNT.

1. Star size. SNT is composed of dots 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 an infinitesimal dark dot is not visible, SNT uses intense small dots ("stars") on a dark background. Provided 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, a property that is important to the scheme.

2. Star positions. The spatial position of each star in the texture is randomly assigned from a two-dimensional uniform distribution across possible locations on the surface, so that two regions of equal area on the surface are equally likely to receive a given star. This prevents pattern matching between patches of the surface, and in conjunction with (3) and (4) below, guarantees random uniform

distribution of dots within the image.

3. Star flux. A countably infinite number of stars are sprinkled onto the surface and their fluxes are distributed as $1/t^2$. As a consequence (see below), 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. As we discuss below, this property is not realizable in practice.

4. Star shape. Stars are not infinitesimal spheres, but rather infinitesimal, uniformly luminous flat disks, lying tangent to the surface. Consequently the amount of light reaching the image from a given star is proportional not only to its flux and the squared reciprocal of its distance from the observer, but also to the cosine of the local slant of the surface (see Figure 1). This property is responsible for the imaging property that the expected number of stars exceeding criterion intensity in the image is invariant to changes in surface slant.

Because a single view of SNT provides no information as to distance, slant, or edge locations, SNT can in theory be utilized to study the perceptual effects of motion parallax or binocular disparity cues in isolation, without the presence of cue conflicts, under the assumption that the visual system appreciates the absence of texture cues when they are, in fact, absent. In practice, the utility of SNT is limited by two factors. First, physical display devices are limited in their ability to portray small, high intensity light sources. Second, using an infinite number of stars is not possible, as this would make the image infinitely luminous, for the SNT.

This rest of the paper is organized as follows: in Section 2, we introduce the

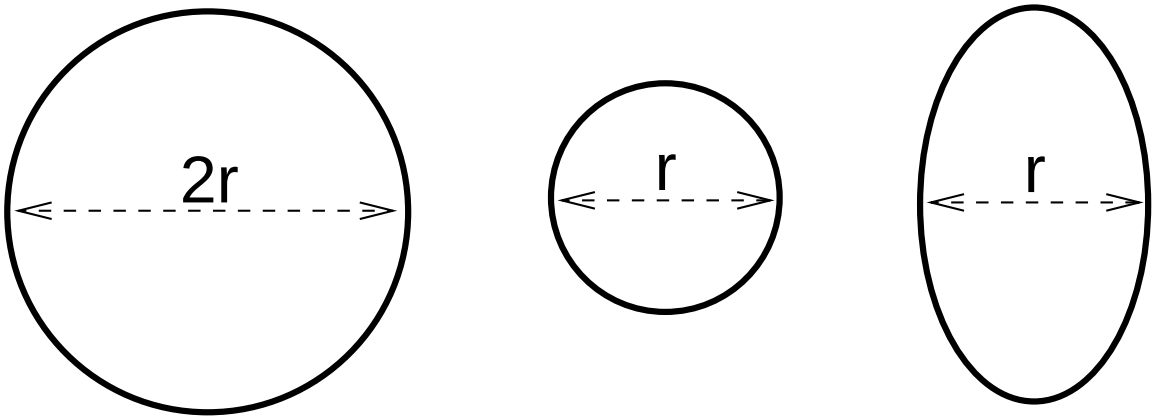

Fig. 1. Idealized images of an infinitesimal luminous flat disk. ¿From left to right: cartoon depictions of the disk's image frontoparallel at a distance of $1\,m$; frontoparallel at a distance of $2\,m$, and slanted by $60°$ at a distance of $1\,m$. Luminance of the imaged area does not change with distance or slant, so luminous intensity is proportional to the disks' image area.

theory of SNT and describe how it can be implemented in practice. We describe the statistics of the texture and give an algorithm for its construction. In Section 3, the claim of invariant appearance across changes of distance is tested experimentally. In two further experiments, the ability to match texture across change in distance is compared for SNT and other textures; and the consequences of averaging across different types of depth cues by the visual system are examined for SNT and other textures. Section 4 is a summary.

2. Theory and Implementation

We begin by defining symbols that will be needed in the paper.

- An "infinitesimal flat star" (IFS) 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

$$L = l \times \cos(\phi)/\pi \tag{1}$$

  candelas (cd), where l is luminous flux in lumens (lm).

- The *Starry Night Texture* in the world consists of randomly arranged IFS's tangent to a surface, with flux values that are distributed as $1/t^2$.

- SNT is fully characterized by two parameters: the lower bound on IFS flux, α, and the *densintensity*, $\mathcal{D}$. $\mathcal{D}$ is the number of IFS's per m^2 expected to have luminous flux $\geq 1\,lm$.

- A *Starry Night Image* (SNI) depicts only SNT-ed surfaces. A SNI can be represented by a set of scaled Dirac vectors $\vec{\delta}(p)$, $p \in \mathbb{R}^2$, where p is the depiction locus of an IFS in the SNI.

- The density of IFS's in the world (number of IFS's per unit of area) is denoted d_W. Their density in the image is denoted d_I.

- The expected number of IFS's, per m^2, per lm, to exhibit a flux value within

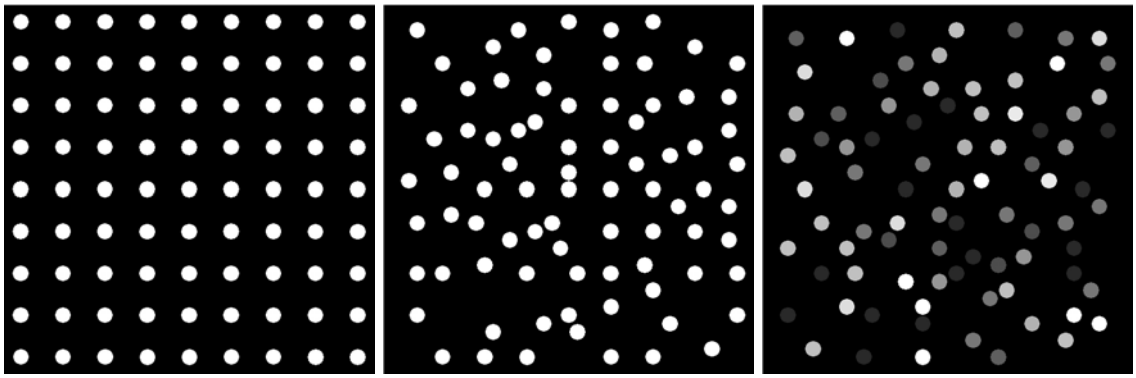

Fig. 2. Cartoon illustrations of textures A, B, and C (left to right). Gray levels represent flux.

the interval $[t-\epsilon, t+\epsilon]\, lm$, where $\epsilon > 0$ and $\epsilon \to 0$, is given by function $D_W(t)$, where $t \in \mathbb{R}_+^*$ and, thus, $\int D_W(t)\, dt = d_W$.

- The expected number of depicted IFS's to exhibit a magnitude of luminous intensity at the direction of surface normal per m^2 of image, per cd, within the interval $[q-\epsilon, q+\epsilon]\, cd$, where $q \in \mathbb{R}_+^*$, $\epsilon > 0$ and $\epsilon \to 0$ is given by function D_I and, thus, $\int D_I(l)\, dl = d_I$.

Three other textures besides SNT are discussed: A, B, and C. In particular: A consists of a lattice of IFS's of constant flux against a black matte background; B is as A, but the locations of the IFS's are uniformly distributed by area; C is as B, but the distribution of flux among the IFS's is uniform (within some finite flux interval). Figure 2 illustrates textures A, B, and C. SNT is shown in Figure 10.

A. *The ideal SNT*

The ideal SNT is composed of infinitely many IFS light sources on a black matte surface. Their locations and intensity values are distributed as follows:

- The locations of the light sources on the matte surface are independent and uniformly distributed by area.
- The distribution of the flux values of the IFS light sources is described by the Probability Density Function (PDF): $D(t) = 1/(kt^2)$, where $t \geq \alpha$, $\alpha \in \mathbb{R}_*^+$, and $k = \int_\alpha^\infty D(t)\,dt$.

The definition above describes stochastically the distribution of flux among the IFS light sources in the SNT. It does not specify the actual number that will occur per unit area.

B. *Absence of distance and slant information in images of SNT*

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 change in distance and slant are a consequence of homogeneity[1] in the texture, which yields a usable depth cue.[5] Of course, knowledge of a particular instance of the SNT does permit an estimate of slant and distance. However, knowledge of the

[1]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.

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 Figure 3.

1. Proof that SNI is devoid of depth cues

For convenience, we assume α to be small, and that SNT can be characterized by $\mathcal{D}$ alone. This assumption is justified as follows. Since $\alpha > 0$, 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. We refer to the luminance of the regions between discriminable IFS's as *background luminance*. Background luminance is a monotonously increasing function of distance and/or slant for realizable SNT. However, Experiment 1 showed that humans are unable to make use of this cue across large changes in distance. Since differences in background luminance were not detectable over some range of distances, we assume that within this range, differences in background luminance are not usable depth cues.

Given that the ideal SNT can be described by a single parameter, its densintensity $\mathcal{D}$, we aim to show that the statistical properties in any region of the image are invariant to the distances and slants of the SNT-textured surfaces that the region depicts. We define an imaging system that obeys the standard laws of perspective projection. To simplify the proof, we will let images be formed on a spherical imaging surface that lies $\rho = 1\,m$ from the center of a projection. For convenience, we also identify the imaging process with the process that replaces an IFS in the world (a "distal" IFS) with an IFS in the image, such that a monocular observer at the center

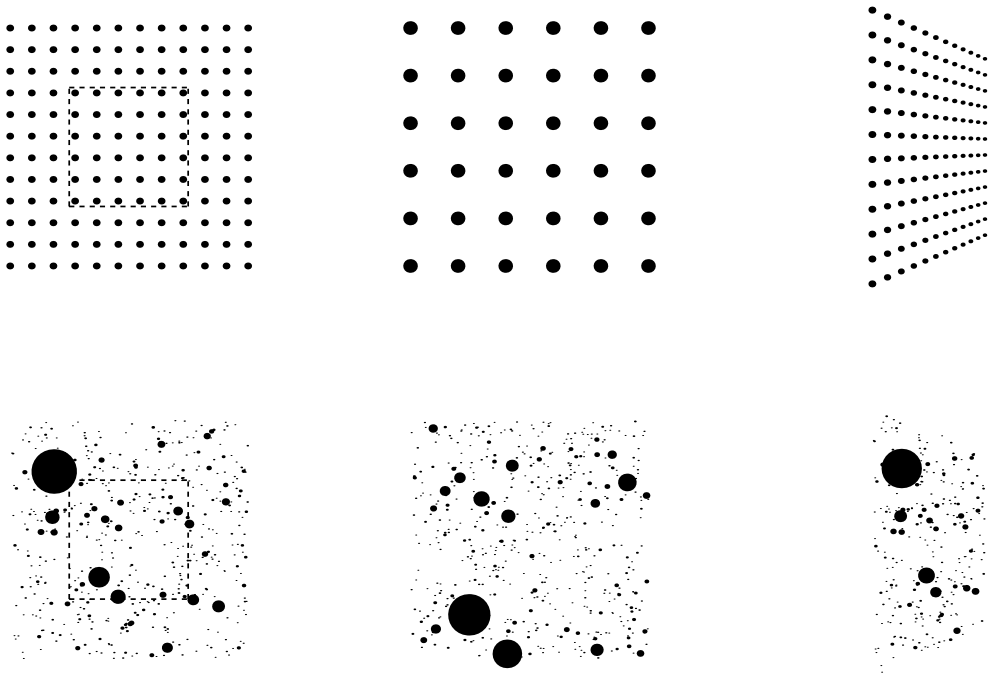

Fig. 3. The strength of depth cues due to texture can be dependent to the statistics of the texture. *Top row:* Common textures are homogeneous and/or isotropic and, thus, their images provide cues to relative distance and/or slant. A homogeneous and isotropic texture (A) on a planar surface patch is imaged from three viewpoints (left to right): (1) frontoparallel at distance L from the patch center, (2) same as (1), but at distance $L/2$, and (3) same as (1), but 60° eccentrically. Circle radii indicate intensities of the imaged IFS's, not the shapes of texture elements. The dashed square in the left panel shows the portion of the surface that is visible in the middle panel. *Bottom row:* When using the SNT the compelling depictions of distance and slant become absent. The panels depict real SNT ($\alpha > 0$) for the same surface and viewpoints as in the top row, in columnwise correspondence.

of projection sees a point light source in the same direction and of the same luminous intensity, whether looking at the distal IFS or at its image; this process is defined in the next paragraph. For the same reason, we also assume (1) that the image is filled by the depictions of SNT-textured surfaces, (2) that the depicted surfaces are smooth and, thus, orientations of the IFS's on the surfaces are well defined, and (3) that there exists an upper bound on the local curvature on those surfaces.

The image is first segmented into a finite number of nonoverlapping openly bounded regions that fill the image, such that any depth edges lie at the boundaries of image regions. Thus, no such region images a depth discontinuity. Next, each of these regions is segmented into a set of nonoverlapping openly bounded subregions, or *windows*, each one small enough to be covered by a disk of diameter ϵ. The IFS's on a surface are countable; e.g. one can count them in order of decreasing flux. Thus, every IFS will be imaged in some window, and the probability is zero that any IFS will be centered exactly on a boundary between windows. At depth discontinuities caused by a slant of $\phi = 90°$, the image intensity of a dot is zero by definition (Equation 1). As the value of ϵ gets smaller, the image within each window approximates a flat image formed by scaled orthographic projection. This approximation is valid because (1) $\rho \gg \epsilon$,[7] (2) the image formation surface is spherical and, thus, everywhere perpendicular to the optical axis,[8] and (3) surface curvature is bounded. Indeed, for any $\delta_1 > 0$, there exists an $\epsilon > 0$ such that δ_1 is an upper bound on the difference in the image positions specified by perspective projection and scaled orthographic projection. Let S_i be the planar patch in the world, imaged by image region i and which minimizes its deviation from surface patch i. Let I_i be the flat image of S_i

formed by the scaled orthographic projection that minimizes image difference relative to the image of S_i in window i. These minimizations can be conceived as the ones that minimize e.g. the sum of squared distances of corresponding IFS's or IFS's images, respectively (illustrated in Figure 4).

We proceed by showing (1) that each I_i is itself textured by SNT, and (2) that the SNT of I_i has densintensity $\mathcal{D}$, the same as the distal surface. This will complete the proof. To establish (1), we must show that the x and y image coordinates of the IFS's in I_i are distributed uniformly and independently, and that their intensities are distributed as $1/t^2$. First, note that the spatial layout of the IFS's in I_i is the same as in S_i after two scaling operations: uniform scaling of linear dimensions by a factor of $1/D_i$, where $D_i\,m$ is the distance to S_i, and scaling by $cos(slant_i)$ in the tilt direction of S_i, where $slant_i$ is the slant of S_i. Thus, the IFS's in I_i have spatial positions that are related to those of the corresponding IFS's in S_i by an affine transformation. Since the x and y coordinates of IFS spatial positions in S_i are distributed independently from uniform distributions, so too are coordinates for the IFS's in I_i[2].

¿From the definitions of SNT and the imaging system, it follows that as ϵ tends to zero the flux of each IFS in I_i becomes equal to the flux of the corresponding IFS in S_i, times the factor $k_i = cos(slant_i)/D_i^2$, where D_i is the distance of IFS i from the

[2]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).

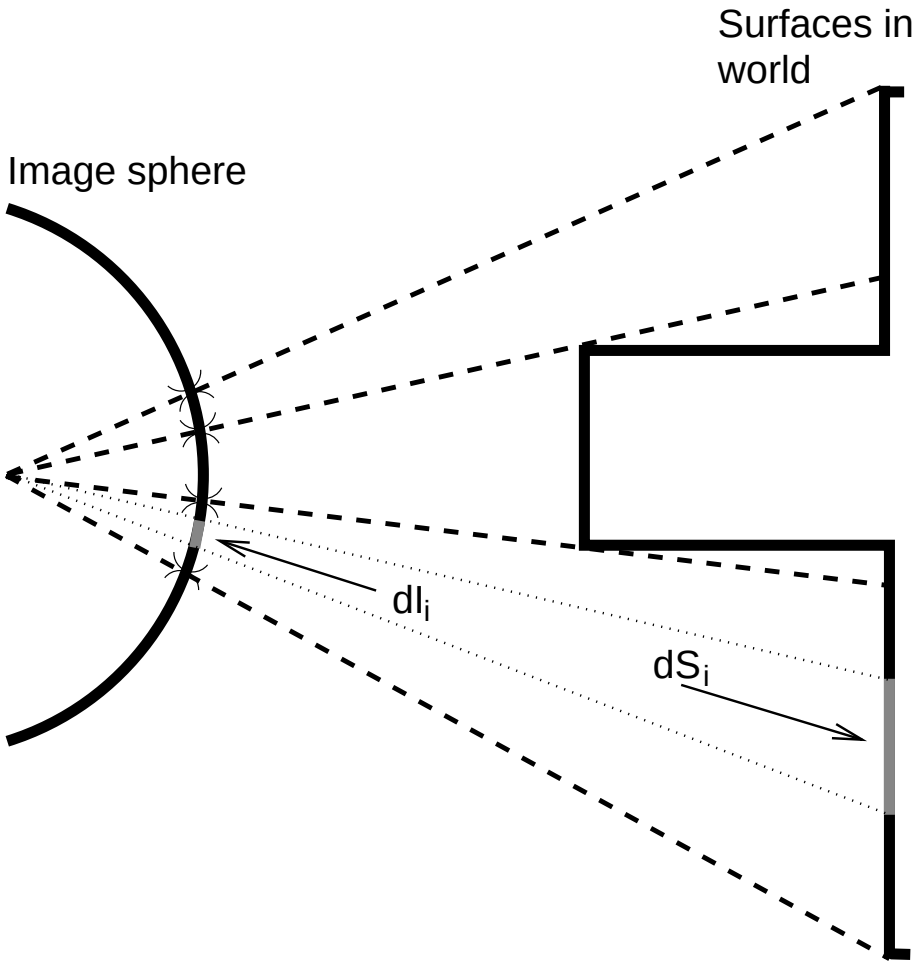

Fig. 4. Illustration of surfaces referenced in the proof of the invariance of the appearance of the SNT to changes of slant and distance (see text). A flat surface patch S_i approximates a surface region. The small and flat surface patch I_i approximates its image.

optical center. Thus the flux of every IFS in I_i is equal to the flux of the corresponding IFS in S_i multiplied by a fixed scalar, k_i. It follows that the fluxes of the IFS's in I_i are distributed as $1/t^2$, so I_i is a SNT.

Finally, if S_i has densintensity $\mathcal{D}$, what is the densintensity, $\mathcal{D}_I$, of I_i? The expected number of IFS's in S_i, per m^2, exhibiting a luminance $1\,lm$ or greater, is $\mathcal{D}$. The area of I_i is k_i times the area of S_i and the flux of each IFS in I_i is k_i times the flux of each corresponding IFS in S_i. Thus, I_i is expected to have $\mathcal{D}$ IFS's of flux k_i or greater, per k_i, m^2. Since I_i is SNT, its IFS's have flux distributed as $\mathcal{D}_I/t^2$. By integrating this quantity, from an arbitrary flux value f_1 to infinity, we see that in I_I, the expected number of IFS's of flux f_1 or greater, per m^2, is $\mathcal{D}_I/f_1$. Thus, two expressions for the expected number of IFS's in I_I of flux k_i or greater, per m^2, are $\mathcal{D}/k_i$ and $\mathcal{D}_I/k_i$. It follows that I_i does indeed have densintensity $\mathcal{D}$.

It can be shown that if the densintensity in two image windows is the same, then the densintensity of the windows' union (together with their shared border if they have one) is the same as in either window. Thus, the image of an SNT-textured scene is itself a SNT-textured scene of the same densintensity.

Intuitively, it is true for any texture composed of IFS's (whether flux is SNT-distributed or not) that if the distance to the surface is halved, only 1/4 of the original surface will remain visible in the window, and the image intensities of the visible IFS's will quadruple. Similarly, if the slant of the surface increases from 0° to 60°, then twice as much surface will be visible in the window, but the image intensities will be halved. We suspect that only when flux is distributed as $1/t^2$ can the distribution of IFS image intensities remain unchanged across such changes in

distance or slant.

Based on the above, and assuming that background luminance provides no usable cue, one can see why a single view of a SNT-textured scene does not provide any clues as to the distances, slants, or edges of surfaces in the scene.

As mentioned above, the ideal SNT is impossible to realize for two reasons:

- The distribution $\mathcal{D}/t^2$ for IFS flux in SNT implies an infinite number of IFS's, since: $\int_\alpha^\infty (\mathcal{D}/t^2)\, dt = \infty$, as $\alpha \to 0$. It is impossible to draw from a fixed positive distribution an infinite number of times, without creating an infinitely luminous image. (It *is* possible to build an infinite set of IFS's that together do not create an infinitely luminous image: the sequence of IFS's having flux 1, 1/2, 1/4, 1/8, ... *lm* is such a set). In astronomy, *Olbers' paradox*[9] is a similar puzzle: if stars were point sources of constant flux, that were distributed with constant density throughout an infinite universe, then the distribution of stars in the night sky would be as $1/t^2$ and the night sky would be infinitely bright.

- There exists an upper limit to the flux of IFS's that can be portrayed by physical displays. Even if this were not the case, an upper bound on flux would be needed to make displays safe to view.

Thus the distribution of IFS fluxes in any realizable SNT must be tailored so that image luminance is not infinite, and also, no star is too intense. We therefore impose

thresholds α and β on the PDF of IFS flux as follows:

$$D(t) = \begin{cases} \tau/t^2 & \text{for } \alpha < t < \beta \\ 0 & \text{for } t \leq \alpha \\ 0 & \text{for } t \geq \beta \end{cases} \tag{2}$$

The term τ in the numerator above is a constant of proportionality that makes the area under the curve equal to 1 and is equal to $(\alpha\beta)/(\beta - \alpha)$. Once n is set, any given star's intensity can be drawn from this distribution. Densintensity now determines the expected (finite) number of IFS's per unit of area, in addition to being equal to the expected number of IFS's per m^2 of intensity greater than $1\,lm$[3].

2. The SNT from a Bayesian perspective

Bayes' Theorem provides a theoretical framework for understanding how the image of a scene can lead an optimal observer to infer the scene that has most likely given rise to an image. To make an optimal decision, the observer must take into account three factors: (a) the prior distribution (frequencies of scene occurrence in the world) and (b) the likelihood that the acquired image could have been created by each scene. The ideal observer also needs (c) a cost function, to map the posterior probabilities of each scene to the choice of a particular scene. (see,[10] for an overview).

Given a SNI, infinite surface arrangements could have given rise to that particular image, since all these arrangements give rise to the (stochastically) same image. In principle, each possible interpretation of the image is equally probable. From the

[3]In principle, the lower threshold α can be greater than $1\,lm$. For completeness, we therefore define the densintensity parameter as specifying the distribution from which the stars are drawn, rather than the actual number of expected stars per m^2 of intensity greater than $1\,lm$.

Bayesian perspective, all possible images exhibit equal likelihood, or otherwise, the resulting image does not convey any information to the ideal observer, as to which stimulus was imaged.

3. Depth cues due to size, shape, and intensity of the IFS image

Besides texture statistics there exist other cues to depth in a SNI. In practice, real light sources cannot be infinitesimal in size and, thus, should they occur at different distances in the world they would be depicted with different sizes in the image. It follows, that the silhouette boundary of the depiction of a flat and circular light source would be a circle only when the source is gaze-normal. Thus, if dots are rendered to exhibit equal size and circular shape, in the image, then the interpretation of image dot size and shape would be that the observed dots are unslanted. We solve this problem in principle by specifying that dots are to be rendered small enough so 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 this case, no cue to the difference in size as well as shape of rendered dots is available. Such an arrangement can be achieved in practice by increasing the distance between the display and the observer; this has the side effect of making the dots look dimmer. Since we were unable to render dots that were simultaneously very small and very bright, we had to compromise, and dot sizes were visible.

Note also that in the model, the intensity of an IFS in the image is related to its distance from the observer by the inverse square law. This approximation is excellent

for small pupils, but is not exactly correct. The IFS intensity function is directionally selective. If an imaging device has a constant, nonzero pupil size, it will gather light from a larger set of emissive directions when it is close to the IFS than when it is far away. For a gaze-normal IFS, doubling the viewing distance would therefore decrease image intensity by slightly less than a factor of 4, because a larger fraction of the gathered rays would be from the IFS's highest intensity directions.

4. Expected mean luminance and appearance of the SNT

The expected value for mean luminance in the image of an ideal SNT can be estimated as a scalar multiplicate of the integral across IFS flux, t, which is: $\int_0^\infty (\mathcal{D}/t^2)t\,dt = \infty$ and, thus, a surface patch equal to a unit area is expected to exhibit infinite intensity. It follows that its image therefore exhibits infinite luminance. Because dim stars are dense in the surface, and therefore in the image, probably any piece of the image will also have infinite expected mean luminance. In the case of SNT thresholded with lower intensity cutoff α, the expected mean luminance is also infinite ($\int_\alpha^\infty (\mathcal{D}/t^2)t\,dt = \infty$). In this case, the possibility of IFS's with very high intensity causes the expected value of mean luminance still to be infinite. However, any given image of a patch of SNT, thresholded in such a way, counts a finite number of dots, each of finite intensity. Thus, the image will have finite luminance. This analysis shows that mean luminance is not a useful way to characterize visual stimuli based on the SNT.

When the viewpoint of image acquisition is very close to a thresholded instantiation of the SNT, all of the dots within the imaging window will be bright enough to see, and will be separated by black image regions. In principle, the lower end of the

distribution ($t \leq \alpha$) could be tailored in other ways besides a cutoff to zero intensity. It is even possible to put an infinite number of dots into this dim portion of the distribution, without causing background intensity to be infinite: for example, by using a $1/t$ or $1/\ln(t)$ distribution function below the cutoff intensity. In this way, the rate at which the foreground gets darker as an observer approaches a (foreground) SNT-ed surface is reduced, and would at least never becomes zero. But still, the brightness of a SNT background of the same densintensity, if visible, would be higher.

5. A numerical example

The effect of distance on the distribution of IFS image intensities is illustrated in Figures 5 and 6. Three textures are considered: B, C, and realized SNT. A histogram of IFS luminous fluxes in the image is shown for an instance of each texture at each of three distances ($2L$, L, $L/2$). Figure 5 shows that for B and C, the IFS's in the image become more closely spaced, and less intense, as viewing distance increases. It is therefore easy to distinguish which of two images represents a closer surface.

Figure 6 shows that the distribution of IFS fluxes in the image is invariant to distance, in the case of SNT. IFS fluxes in the image remain distributed as $1/t^2$ and the distribution does not shift up or down. The figure also shows the limitation of the realized version of SNT: the SNI gains additional dim stars, and loses bright stars, as the distance of the surface in the world is increased.

For these three textures, a change in slant affects the distribution of image fluxes in a manner similar to, but not identical to, a change in distance. As slant increases, both B and C realize an increase in IFS density and decrease in flux, as with distance.

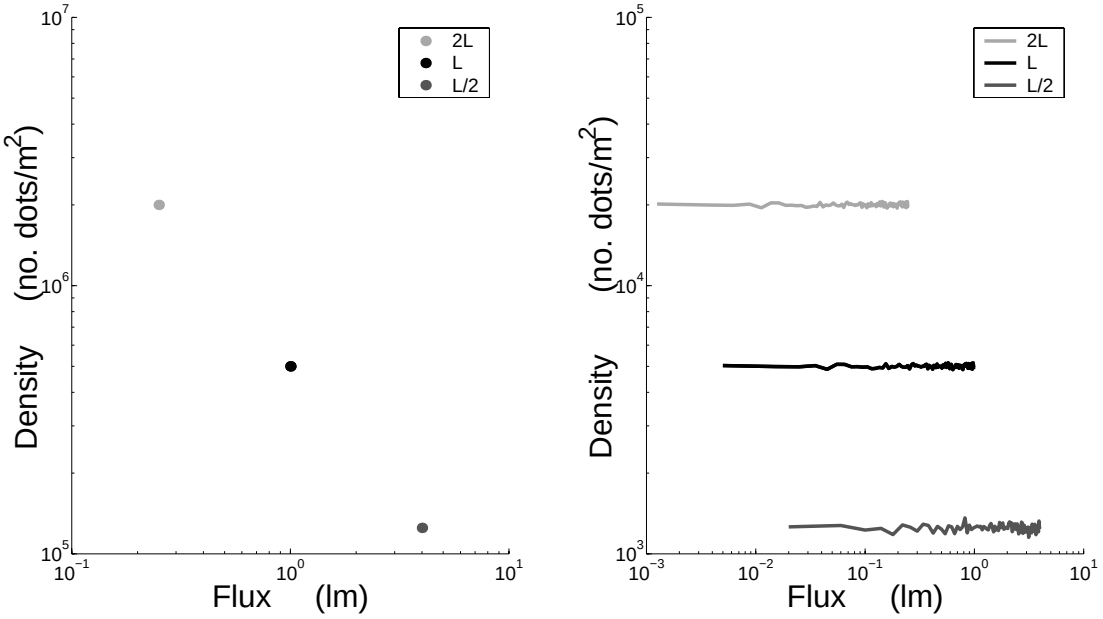

Fig. 5. Histograms of IFS image flux for textures B (left) and C (right), for frontoparallel surfaces at three distances ($2L$, L, $L/2$). The abscissa plots luminous flux and the ordinate counts the number of IFS's per unit area in the image at a given flux. Both axes are logarithmic. Bin width was fixed in linear units. For texture B, all IFS's appear at a single value of flux that depends on distance (left to right: $2L$, L, $L/2$). For C, fluxes cover a range of values (top-left to bottom-right: $2L$, L, $L/2$). As distance increases, the maximum IFS flux in the image decreases, and there is an increase in the density of IFS's in the image at represented IFS flux values.)

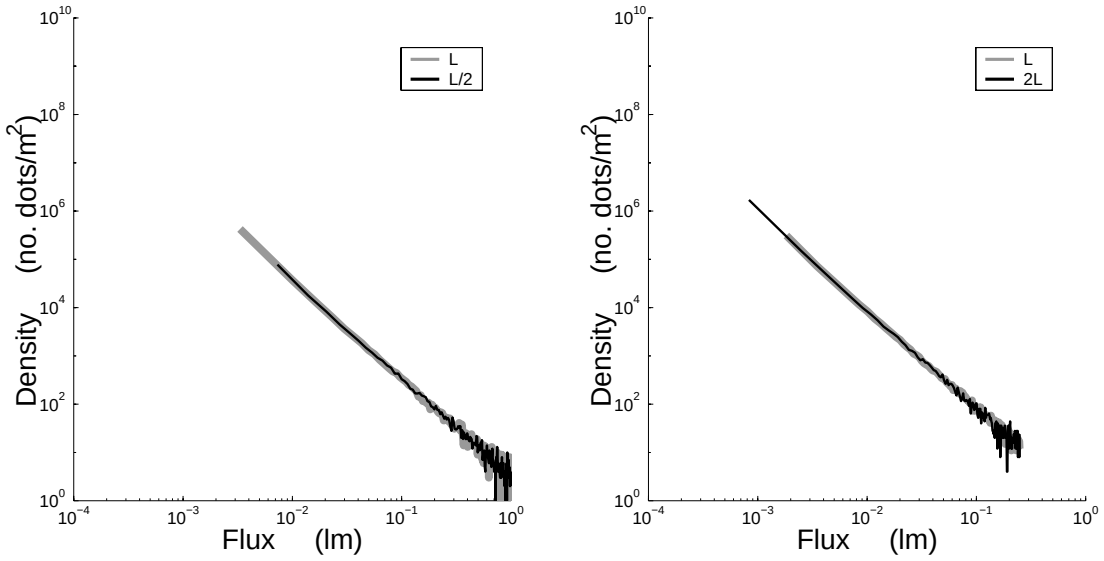

Fig. 6. Histograms of IFS flux in the image for SNT, for frontoparallel surfaces at different distances. Axes are as in Figure 5. Left: Histograms for a thresholded (realizable) version SNT at distances L(bold gray) and $L/2$ (black). Right: Distances L (bold gray) and $2L$ (black).

Under perspective projection, slant also has a secondary effect, as it causes some distal IFS's to become closer, and others farther, from the observer. As with changes in distance, changes in slant cause dramatic qualitative changes in the flux histogram for B and C, and no such change for SNT, except at the ends of the flux distribution.

6. A physical approximation of SNT

Artificial gold leaf provides a crude approximation to SNT. This material is extremely thin. It is opaque (unlike gold leaf, which appears green when lit from behind) and versions can be obtained that have a great many very small holes of varied size. When backlit, the holes appear circular and the smallest holes are not individually visible. The distribution of the sizes of these openings (and thus their luminous intensities) of the material was not analyzed, but it is similar to the one required for SNT in at least one respect: large holes occur and are rare. Figure 7 shows a device we built to demonstrate properties of SNT. Two pieces of artificial gold leaf ("Sepp Guilding Workshop Imitation Gold Leaf") were lit from behind using two light boxes, that were mounted on tracks so that the surfaces could be adjusted in distance or slant. Observers reported being unable to see factor of two changes in distance, and 40° changes in slant, in monocular views of the surfaces.

C. Stimulus Formation

We now present an algorithm for generating a SNI, taking into account imaging limitations of the visualization medium. This algorithm and its variants were used to generate the SNI's in the modeling examples and experiments.

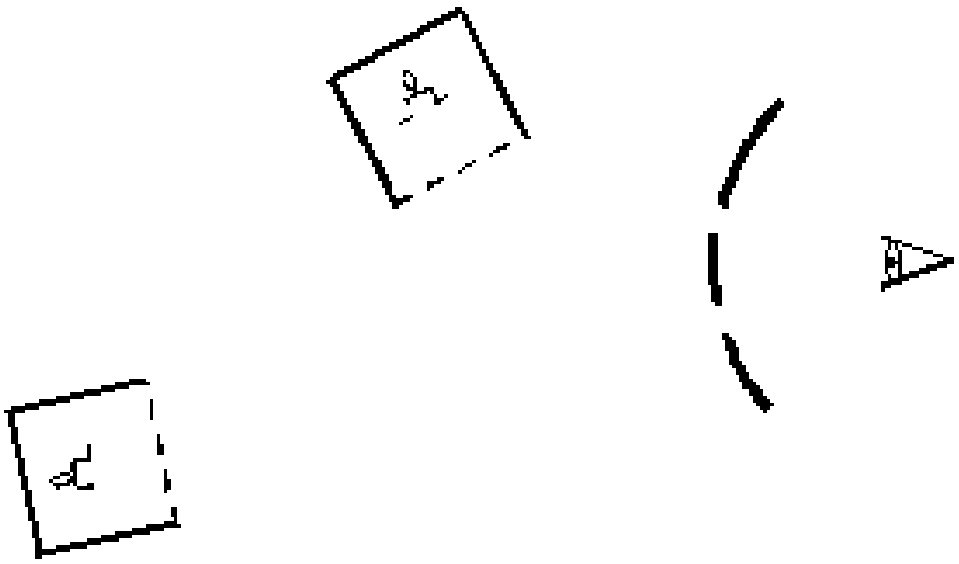

Fig. 7. A device to illustrate SNT's ability to hide distance and slant information. The two boxes can be adjusted to different distances and/or slants. Each box is viewed through a tube that creates a fixed aperture and restricts viewing to a central portion of the box face. The inside of each box is lined with reflective aluminum foil. The front face (dashed line) is translucent plastic, to which a piece of artificial gold leaf has been applied. A lamp illuminates the leaf from behind, and a black curtain blocks direct view of the surfaces.

The process of physically realizing the SN stimulus, or otherwise generating a SNI, has the following steps: (a) modeling the surfaces to be rendered, (b) instantiating the texture on the surfaces, (c) generating an abstract representation of the image (IFS locations and fluxes in the image), and (d) rendering the final image. In Figure 8, the flowchart of this process is shown.

1. Modeling of surfaces

In this step of the algorithm, the model of the three-dimensional scene to be rendered is determined, by establishing the positions of points on the surfaces of the depicted scene. Data structure M contains the 3D coordinates of points on the visible regions of the surfaces. A local surface slant is associated with each point in M, either explicitly, or implicitly by virtue of the point's belonging to a particular surface; information at this stage must be sufficient to compute the distance and slant of each depicted surface point, relative to the observer. Distance and slant are required for the computation of the luminous fluxes of the imaged IFS light sources, in step (c) of the algorithm (see Section 4).

2. Instantiation of the texture

The second step of the algorithm is the instantiation of the texture, that is, to determine the number, location, and flux values of the imaged IFS light sources. The input for this task is the density, the type of texture, and data structure M. A virtual, calibrated, camera that will image the scene is also assumed.

The number n of the represented light sources is determined, based on texture density and the area of the represented surfaces in M. The generation of the cor-

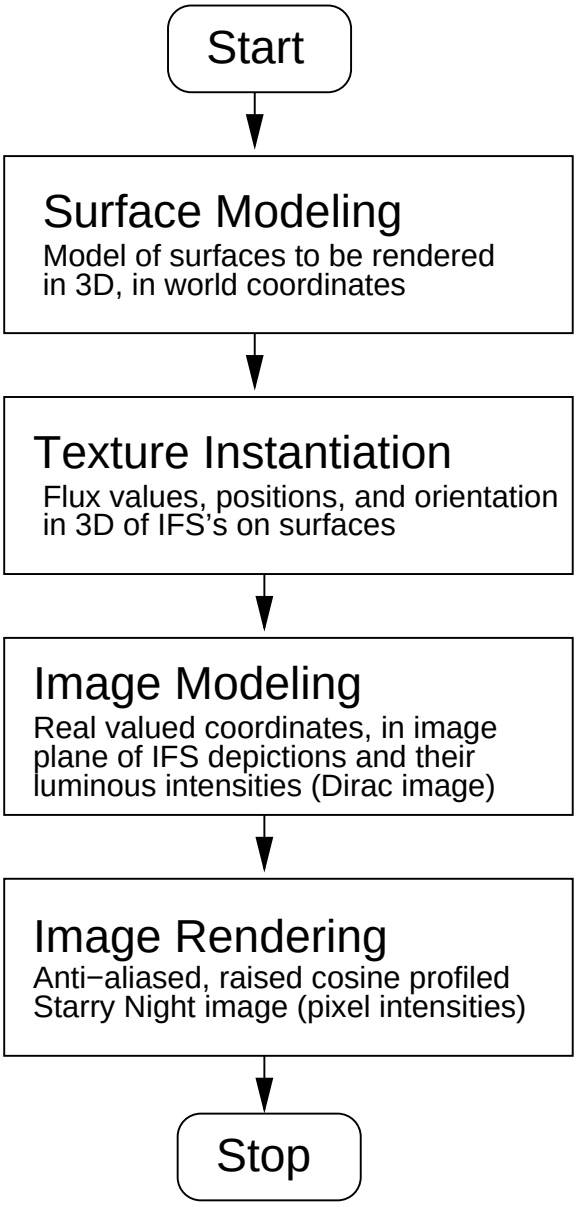

Fig. 8. The flowchart of the SNI formation algorithm.

responding n flux values requires as input the type of texture to be rendered. For textures A and B, the flux values are constant. For textures C and SNT, a pseudo-random number generator is utilized. Upon invocation this generator returns value $R \in [0, 1]$, where each possible value of R is of equal probability. Given the PDF $D(t)$ of the flux values of texture C, where $t \in I = [\alpha, \beta]$, the distribution can be instantiated by computing n times the function $(\beta - \alpha) \times R + \alpha$. For the SNT, the problem may be formulated as follows: given Equation 2, for which $\int^I D(t)\, dt = 1$, distribute a set of numbers in I so that the probability of values occurring in some interval $[v, v + \epsilon] \subset I$ is $\int_v^{v+\epsilon} D(t)\, dt$. But since, $\int_v^{v+\epsilon} D(t)\, dt = \epsilon/(v(v + \epsilon))$, a value drawn from the distribution can be generated as $\epsilon/(1 - U\epsilon)$, where $U = R \times ((1 - \epsilon^2)/\epsilon) + \alpha$.

The locations of the depicted light sources are random and of equal probability. The corresponding, n, 3D-coordinates can be generated by assigning $2n$ pseudo-random real numbers to the x and y coordinates to form planar textured patches; the z-coordinates are set to zero. Then translation and rotation matrices can be employed to apply these patches on the depicted surfaces, so that they cover them completely. Intuitively, the process can be described as the construction of textured "wallpaper" pieces, which are then applied on the surfaces of the scene. The ranges of the pseudo-random coordinates are determined by the size of the imaged surfaces.

The completion of this step of the algorithm yields structured data that represent the locations of the luminous sources in space and their flux values.

3. Creation of the image model

In this step of the algorithm, a model of the image to be rendered is generated. For this purpose, the optical rays that reach the camera center are considered and their intensity values in the image calculated.

In order to detect which light sources are not visible due to occlusions, the imaged surfaces are perspectively backprojected from the camera center, until another surface is encountered. The light sources that are encountered in this process are discarded.

In order to calculate the flux values of IFS depictions in the image, the scene structure, represented in data structure M, is considered. The luminous flux that a depiction of an IFS is required to exhibit, in order to simulate an IFS in the world, is:

$$l_I = l_W \times \cos\phi \times \left(\frac{r_I}{r_W}\right)^2 \times \frac{1}{\cos\theta} = l_W \times \cos\phi \times \left(\frac{r_S}{r_W}\right)^2 \times \frac{1}{\cos^3\theta}, \qquad (3)$$

where l_W is the flux of IFS in the world, l_I the flux of IFS in the image, r_W radial distance from observer to IFS in the world, r_S the radial distance from observer to screen, r_I the radial distance from observer to IFS in the image, ϕ the slant of IFS in the world relative to line of sight, and θ the eccentricity of both IFS's relative to screen normal. Figure 9 illustrates the above.

The outcome of this step of the algorithm is a data structure, henceforth referred to as a scaled Dirac image SNI_δ. This data structure represents the locations of the images of the depicted light sources and their luminance values, as Dirac functions that are centered at some point on the image plane and multiplied with a scalar. The value of this scalar represents the magnitude of luminous intensity of the image of

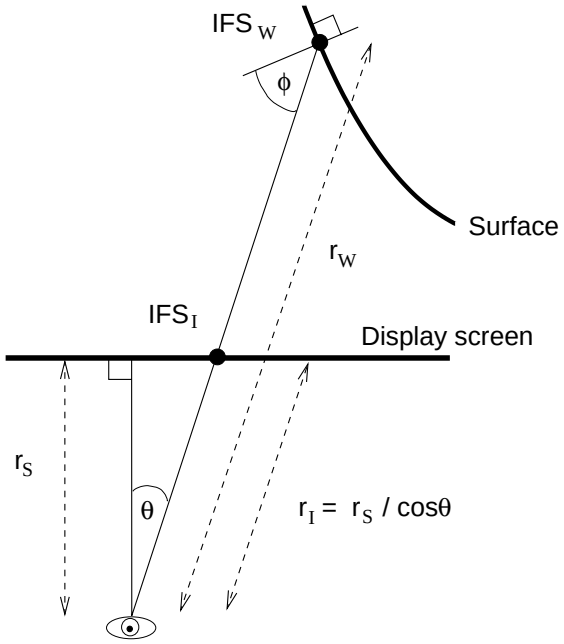

Fig. 9. The geometry of simulating a distal IFS (IFS_W) in the world with a luminous dot on a screen (IFS_I).

the depicted IFS light source. The center of the Dirac impulse represents its location. Data structure SNI_δ is treated as a list of 2D coordinates instead of as an image, thus conserving the precision of the IFS depiction coordinates.

4. Rendering

In this step of the algorithm, the rendering of the SNI is performed, based on the SNI_δ data structure. In particular, the tasks accomplished are: the localization, the rendering of the depictions of the imaged light sources, and a final simple filtering of the image.

The locations of the light source depictions are encoded in SNI_δ, as floating point numbers. An anti-aliasing procedure can be used to facilitate dot placement at arbitrary positions between integral locations, for a discrete coordinate display (e.g. a CRT). This procedure increases the localization accuracy with which dots are rendered on the SNI and is based on 2D, bilinear in our implementation, interpolation. It ought to be noted that since the coordinates of luminous dots on the SNI are treated as independent variables, when two luminous dots occur at the same locus their luminance values are accumulated at that image point. Given the anti-aliasing procedure described above, this accumulative process is performed for dots which, although may not coincide, are proximate enough for their anti-aliased representations to overlap. Thus, the anti-aliased representation for each dot is computed independently and the results are additively summarized in the final image. The result of the process described above is an image, which is henceforth referred to as anti-aliased scaled Dirac image.

The way that depictions of light sources are rendered is determined by their desired size, in the image. In the case of the smallest size ($2 \times 2\, pixels$), the luminous dots are rendered in the SNI as outputted from the anti-aliasing procedure, described above. If a larger size is used, then the spatial distribution of luminance in the image was chosen to exhibit a 2D raised cosine pattern as: $(1 + \cos d)/2$, for $d \leq \rho$ and 0 elsewhere, where ρ is the radius of the kernel and d the distance from its center, for the following reasons. First, its the pattern is rotationally invariant and, thus, compatible with the mental model of a ray of light intersecting the image plane. Second, the raised cosine kernel gathers most of the luminous energy near the locus of the center of depiction of the IFS light source. Third, all of the energy is enveloped within this kernel (e.g. as opposed to a Gaussian kernel). After normalizing kernel values so that their sum is 1, the final image can be then rendered as the result of the convolution of the anti-aliased scaled Dirac image with the raised cosine kernel.

Finally, to compensate for the typically poor linearization of CRTs' nearby the lower end of its dynamic range, a 5% of the CRT's intensity range is added to each image pixel. IFS flux values are chosen such that no overflow occurs.

5. Result (computer-generated SNT stimulus)

A sample output of the algorithm is shown, in Figure 10. Note, however, that the images of SNT are probably not displayed veridically on this paper, since we did not compensate for the gamma function of these displays. For this reason in Figure 10, the logarithm of each pixel value was computed and the resulting image is shown on the right.

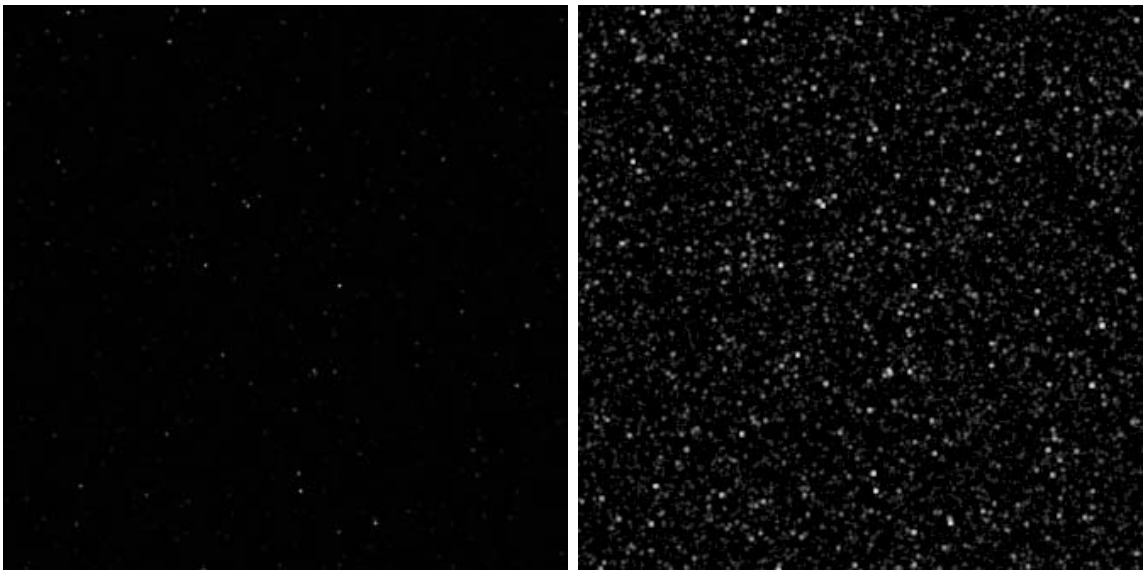

Fig. 10. A Starry Night Image. Left: the original image, with no correction of the linearity of image intensity values, for the reflectance of the paper. Right: The same image, with its intensity values logarithmized.

3. Experiments

Three experiments were performed to characterize specific aspects of human perception for images of SNT. First, it is important to determine the perceptual consequences of truncating the flux distribution for IFS's in the SNT. Experiment 1 therefore compared distance-discrimination thresholds for simulated surfaces textured with SNT to thresholds for three other IFS textures (textures A, B, and C). To anticipate the result: for the SNT we used, the cutoffs caused no visible differences between the images of near and far SNT surfaces, for 100-fold changes in distance; in contrast, a 10% 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 the densities of texture elements on surfaces at different distances are the same, or if one texture is denser than another in the world. If observers can make estimates of this sort for SNT, they ought to adjust the densintensity of SNI to be the same for surfaces depicted at different distances, when trying to match textures in the world. In Experiment 2 we asked observers to match textures at different distances, and we found large effects of simulated distance for textures B and C, but not for SNT. Third, our interest in SNT derives from its ability to hide information about distance and slant, so that nonpictorial depth cues can be studied in isolation. Is it the case that our SNI differed measurably 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 and texture cues. In the stimuli, for both SNT and texture B the binocular disparity cues indicated nonzero slant. For the SNT the texture cues were uninformative, whereas for texture B, texture indicated zero slant. We found that slant was underestimated for both SNT and texture B to the same extent. This suggests that accommodation by the lens of the eye, or other depth cues, were 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 large monochrome 19 inch *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(s). 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 $6 \times 10 cm$, rounded in shape to fit close to the eye, oriented at 45° relative to the line of sight, and positioned $1.25\, m$ above the floor. The face of each CRT was always perpendicular to the line of sight from the eye to the center of the screen. The room was completely dark except for scattered light from the displays, and for Experiment 3, black paper frames with $4\, cm$ saw-tooth edges occluded the edges of the screen to insure against binocular matching of the display edges. A computer generated the stimuli before the experiment, and displayed them during the experiment, using Matlab and the Psychophysics Toolbox.[11] Observer's responses were made using a numeric keypad and collected by the computer.

The aspect ratio of each display was adjusted to 1. The visual locations of the dots in the displays were specified to within ≈ 30 seconds of arc. This high level of spatial precision was achieved by use of two procedures: anti-aliasing and spatial calibration. The anti-aliasing procedure was described in Section 4. Spatial calibration involved the creation of a lookup table that converted desired visual directions into screen coordinates. For the calibration procedure, thin nylon filament was stretched across a precisely milled loom, creating a $36 \times 26\, cm^2$ grid with $1\, cm$ spacing. The loom was mounted $1\, cm$ in front of each CRT in turn, and was viewed in the mirror from the standard viewing position. Dots on the screen were positioned to be coincident with their corresponding intersections in the grid. After adjusting approximately 70 dots, the interpolated positions of the remaining dots were also correct to within the limits of the experimenter's acuity. The loom was then removed, and its former location defined a virtual plane onto which stimuli were projected. The calibrated area was $48^\circ \times 36^\circ$ on each monitor.

The luminance of the CRT's was calibrated to be a linear function of the nominal image intensity value, on a scale of 0 to 255. A photometer was used to measure 16 of the CRT's luminance values, equidistantly distributed across the dynamic range. At each measurement the total of image pixels were set to the same value and the photometer was targeted at the center of the screen. At low levels of luminance where the photometer was insensitive, calibration images were used to linearize the response. Intermediate values were estimated using cubic splines. The two CRT's were calibrated to be equiluminant with each other.

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 to the corresponding ranges for textures A, B, and C. Each texture was tested at several densities (or densintensities for SNT) to ensure that performance was not highly sensitive to density at the values tested.

1. *Stimuli*

Stimuli depicted a frontoparallel background surface. Half of the stimuli depicted a foreground surface as well, that occluded the central region of the background surface. The foreground and background had the same texture (A, B, C, or SNT). Viewing was monocular.

An illustration of the depicted scene's geometry is shown in Figure 11. The background surface was a $44 \times 44\,cm^2$ frontoparallel, flat, square patch at a simulated distance of $r_{bg} = 1\,m$ from the observer. In one of the two stimuli presented in a given trial, the foreground surface was also present. This surface was also frontoparallel and located at a variable distance $r_{fg} \leq r_{bg}$, closer than the background surface, thus partially occluding it. The foreground patch was square and its area in the the world was adjusted as a function of distance so that, in the image, it occluded the central 1/4 of the area of the background square. Both surfaces were centered on the screen. For texture A, 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. Dot size was $4 \times 4\,pixels$.

The texture density range was selected as follows: the lowest density (or densin-

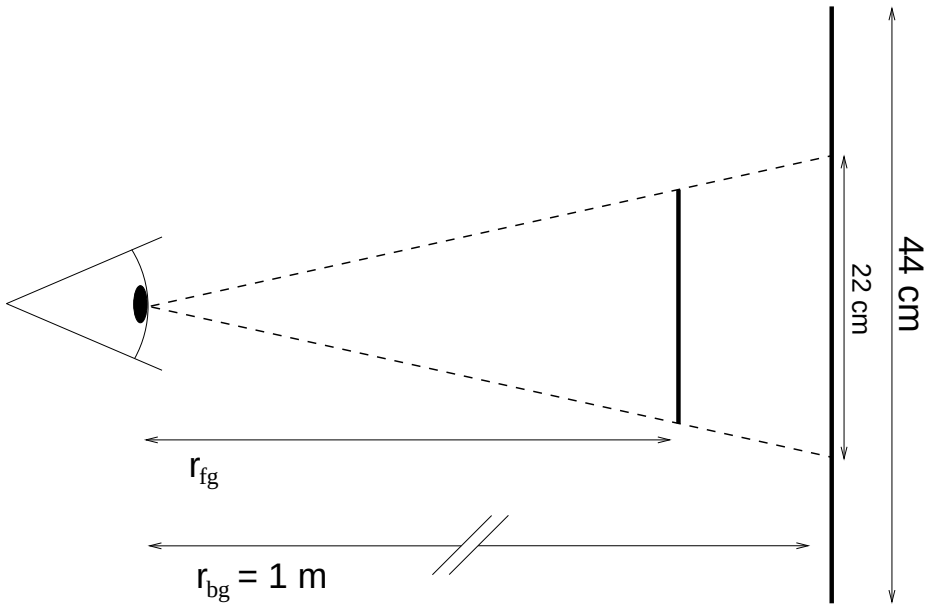

Fig. 11. An instance of the virtual scene depicted by stimuli of Experiment 1.

tensity) was the one for which the texture was just dense enough to do the task. At this density, just a few dots were visible in the image. The highest density was determined by the resolution of the display; it was a density at which that almost all pixels in the image were occupied by some part of a luminous dot. The range of densities d_W was $16\,dots/m^2$ to $625\,dots/m^2$ for textures A, B, and C and $11\,dots/m^2$ to $4000\,dots/m^2$ for the SNT (for SNT, d_W refers to the actual number of IFS's placed onto the surface, per unit area of texture in the world, whether IFS's were visible in the image or not). Finally, the flux range of IFS's was chosen. For A and B, flux was $7.5 \times 10^{-6}\,lm$. For textures C and SNT, flux was tuned separately for each density, such that a maximal number of IFS's were visible, as follows: $3.7 \times 10^{-9}\,lm$ to

$7.5 \times 10^{-6}\, lm$, for C, and $\alpha = 3.7 \times 10^{-9}\, lm$ to $\beta = 4.5 \times 10^{-5}\, lm$ for SNT[4].

2. Task

A two-interval forced choice task (2IFC) was utilized. On each trial, two stimuli were presented for $2\, sec$ each. One of them contained the foreground patch in addition to the background patch; the observer's task was to indicate which interval contained the foreground patch, by making a keypress. Visual feedback was given after each trial. A blank (black) screen was displayed for $1\, sec$ between visual display. A 3-correct, 1-wrong staircase procedure controlled the depth interval between the background and foreground patch. Novel instances of textures B, C, and SNT were used on each presentation. Increments in depth were spaced logarithmically (base 2). The staircase terminated after 15 reversals; threshold was the average of the last 12 reversals. This yielded an estimation of the detection threshold for which approximately 79% of the observer judgments were correct. Each such estimation was considered as one independent measurement. Observers were encouraged to rest whenever in need.

3. Results

Figure 12 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

[4]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) IFS's had this level of flux, due to the sparseness of high-flux IFS's in SNT. In rare cases where 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. This rejection was performed in Experiments 2 and 3 as well.

fractional change in distance between the background and foreground, required for the stimulus containing a foreground and background patch to be discriminated from the stimulus containing only the background patch. Fractional change was computed as $r_{bg}/r_{fg} - 1$ 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-point texture (texture A); a change in distance of approximately 1% was sufficient to identify the foreground-containing stimulus. For textures B and C, roughly 10% change in distance was required. For SNT, a roughly 100-fold change in distance was required.

Observer XZ collected data for IFS densities in the images that ranged from very sparse (very few IFS's in the image) to very dense (very little black space). Across this range, increasing density caused a decrease in threshold; this effect was most pronounced for textures A 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 to see a change in the image, was three orders of magnitude weaker for texture SNT than for textures B or C, and four orders of magnitude weaker than for texture A. We conclude that depth is more difficult to see in scenes that are textured with SNT than other textures.

It is important to remember that the stimuli used in these experiments were not real surfaces, but rather, surfaces rendered on a computer screen. IFS's were simulated using dots of small but visibly nonzero size, and accommodation did not

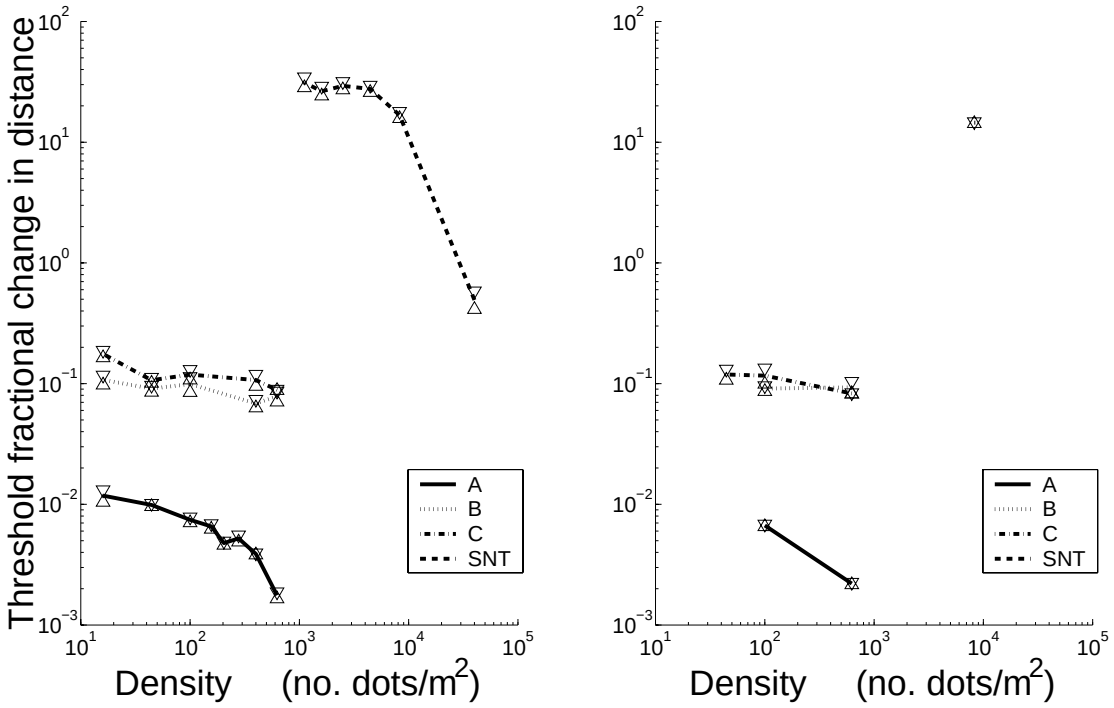

Fig. 12. Data from Experiment 1, for two observers XZ (left) and BTB (right). The graph plots the threshold fractional change in distance required to identify which of two stimuli contained an occluding surface patch, for 4 textures. The abscissa is dot density and the ordinate is fractional change in distance. Data points are the average of 3 threshold measurements from three separate staircase procedures. The centers of the triangles mark the standard error of the mean.

indicate different distances for the foreground and surface patches, as would have been the case for real surfaces. Accommodation (or blur cues) therefore conflicted with the simulated changes in distance, that were specified by texture cues. Do our findings apply to real SNT surfaces? It seems likely that a 100-fold change in distance would indeed be required to see the change in distance for a real SNT surface, but only if it was far enough from the observer to make accommodation and blur cues unusable.

Dot size was another conflicting cue: in our stimuli, dot size in the image did not increase with the decrease in simulated distance. If the visual system is obliged to take dot size into account when judging surface distance, then this countercue might have artificially elevated the distance threshold for SNT. However, the control conditions (texture A, B, and C) show that constant dot size in the image does not generally cause a failure in the ability to use other aspects of the texture cue (such as dot density) for distance discrimination, so it is reasonable to infer that the extraordinary ability of SNT to hide changes in distance did not derive from this artifactual property of our displays.

An aim of this experiment was to determine whether the lower and upper cut-offs for IFS flux, that were needed to realize SNT, would cause failure of the theoretical ability of ideal SNT to hide changes in distance. The cut-offs did not greatly compromise this feature of ideal SNT.

After the experiments, observers reported on the visual cues that they were aware of using to do the task. For textures A, B and C, observers felt discrimination to be based on the fact that the foreground surface appeared sparser and brighter than the background. As texture density increased an additional cue became available in

texture A: 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 inter-IFS regions: when the foreground surface was close enough to the observer, it began 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. SNT also provided another potential cue: the distribution of IFS was clipped from above, so the foreground patch contained, statistically, more of the most intense IFS's than did the background. However, the observers were not aware of using this cue.

It is interesting to note that texture C did little better than texture B at hiding changes in distance. In texture C, IFS fluxes were randomly chosen from a uniform interval that spanned most of the CRT's dynamic range. Randomizing intensity was not sufficient to hide changes in distance; the particular $1/t^2$ distribution of IFS fluxes in SNT is therefore a critical factor in its ability to hide changes in distance. The small difference in threshold for textures B and C results from some combination of two factors: the randomization of flux values in C, and the fact that the dimmest IFS's in C were not actually visible, which lowered the effective density in that texture.

C. Experiment 2

Unlike other textures, the image statistics of an SNT surface do not change when the distance to the surface changes. It is therefore of interest to ask, for SNT, whether

observers perceive a change in the texture in the world, when the distance to the surface changes but the regional image statistics (densintensity of the SNI) do not. In Experiment 2 we measured this indirectly, using a nulling paradigm, in which observers adjusted the density (or 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. We predicted that SNT's would be adjusted until the densintensity of the test surface matched that of the reference surface, independent of the simulated distance of the test surface. At that point, the images of the test and reference surfaces would be very similar, because of the statistical properties of SNT. As a control, we repeated the experiment using textures B and C. For these textures, we predicted that observers would match texture density in the world. Since observers have to use the images to do the task, this prediction implies that observers will set the IFS density in the images to vary as a function of distance.

1. Stimuli

Viewing was monocular. Stimuli depicted two square surface patches of equal size, the *test surface* and the *reference surface*. Both were facing the observer and resting on a checkerboard floor. The reference surface occurred always in the same position, while the test surface varied in distance from the observer. In all stimuli, both surfaces were textured with the same type of texture, one of B, C, or SNT. Scene geometry and a sample stimulus are shown in Figure 13.

The reference surface was positioned at a depth of $20m$ from the observer, on the

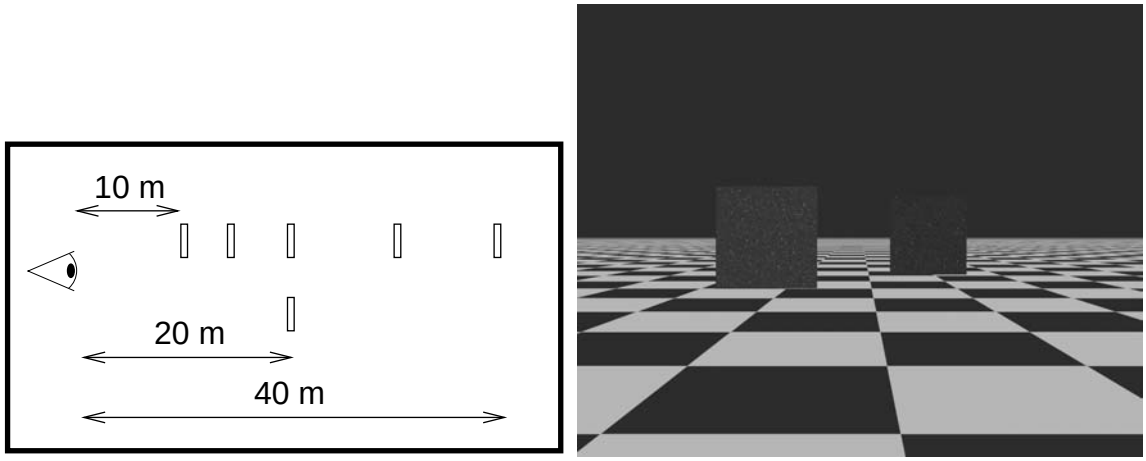

Fig. 13. Left: Schematic diagram of scenes used in Experiment 2 (top view). The eye is shown on the left. The reference surface was at a fixed distance 20 m from the observer, and the test surface varied in distance from 10 m to 40 m; some of those positions are shown in the figure. Right: Sample stimulus as seen by the observer during the experiment. The test surface is at a simulated distance of 15 m.

right side of straight ahead. The test surface was on the left and its distance took on one of 21 values logarithmically spaced between $10\,m$ and $40\,m$. Thus, the distance for the i^{th} test surface was 10×1.0718^i, $i \in \{0, 1, 2, \ldots, 20\}\,m$.

The size of the surfaces, in the world, was $2.2\,m \times 2.2\,m$. All of the surface was visible. The observer was $1.25m$ above the checkerboard floor, which extended nearly to a horizon at eye level. The checks in the checkerboard floor were gray (luminance $42\,lm/m^2$ in the image) and black. The sky was black.

Each IFS on the surface was portrayed by an antialiased dot occupying $4\,pixels$ on the display. The upper cutoff of the flux distribution (β) was selected in the same way as in Experiment 1. On the reference patch, densities for the three textures (B, C, SNT) were selected so that these textures would appear similar in density. For texture B, d_W was $240\,dots/m^2$ and flux $2.6 \times 10^{-3}\,lm$. For C, d_W was $240\,dots/m^2$ and flux values were from $5.2 \times 10^{-6}\,lm$ to $3.2 \times 10^{-3}\,lm$. For SNT, d_W was $3845\,dots/m^2$, $\alpha = 5.2 \times 10^{-6}$, and $\beta = 31.2 \times 10^{-3}\,lm$. For the test patch, density, and in one condition also flux, were adjusted by the observer as the dependent variable(s).

In these stimuli, the cues available to the observer to indicate the ratio of distances to the test and reference surfaces were: the texture gradient and linear perspective in the checkerboard floor, the relative sizes of the surfaces in the image (under the assumption they were the same size in the world), and the declinations in the lines of sight to the bottoms of the surfaces. For conditions B and C, IFS flux in the image was an additional cue.

2. Task

Observers adjusted the density and flux of the texture on the test surface using keypresses, until it appeared to be the same as the texture on the reference surface. Observers were explicitly instructed to equalize these textural properties 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 five conditions, blocked by session. In each condition, one type of texture was used. The conditions were: B, C, SNT, $B2$, and $B3$ using textures B, C, SNT, B, and B respectively. In the first three conditions (B, C, and SNT), the observer adjusted density d_W (and consequently d_I). The distribution of IFS fluxes was always the same for both surfaces in these conditions. In other words, the fluxes of the IFS's in the image were clamped at their correct values, given the simulated distance. In the last two conditions ($B2$ and $B3$), observers adjusted both density and flux.

In conditions B, C, and SNT, a pair of keys was used to adjust the density of the test surface texture. In condition $B2$ an additional pair of keys was available to the observer for adjusting flux, and the observer could alternate at will between adjusting density and adjusting flux. Finally, in condition $B3$, one pair of keys used to adjust density and flux. In this condition, a keypress would increase or decrease both density and flux.

Keypresses caused small changes in density or flux. Density and flux were logarithmically distributed throughout the ranges of allowed values. The range(s) on a

given trial were centered at the density or flux specified by the observer's response on the previous trial, except on the first trial, where the center was pseudo-randomly chosen. 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, DMB; the first two were naive to the experimental hypotheses and the last was one of the authors.

Figure 14 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. The central horizontal line plots the density of the texture on the reference surface, indicating the response predicted by complete perceptual constancy. The diagonal line indicates the predicted settings for complete failure of constancy by matching density in the images. The dashed lines show the limits of the range of densities that was available to the observers.

The data for condition $B2$ are shown in Figure 15. Each column represents the responses of one observer. The graphs in the top row are in the same format as Figure 14. In the bottom row of graphs, the abscissa is distance and the ordinate is IFS flux. The diagonal, horizontal, and dashed lines have the same meaning as in the top row, but now refer to flux instead of density. In all graphs, each data point shows

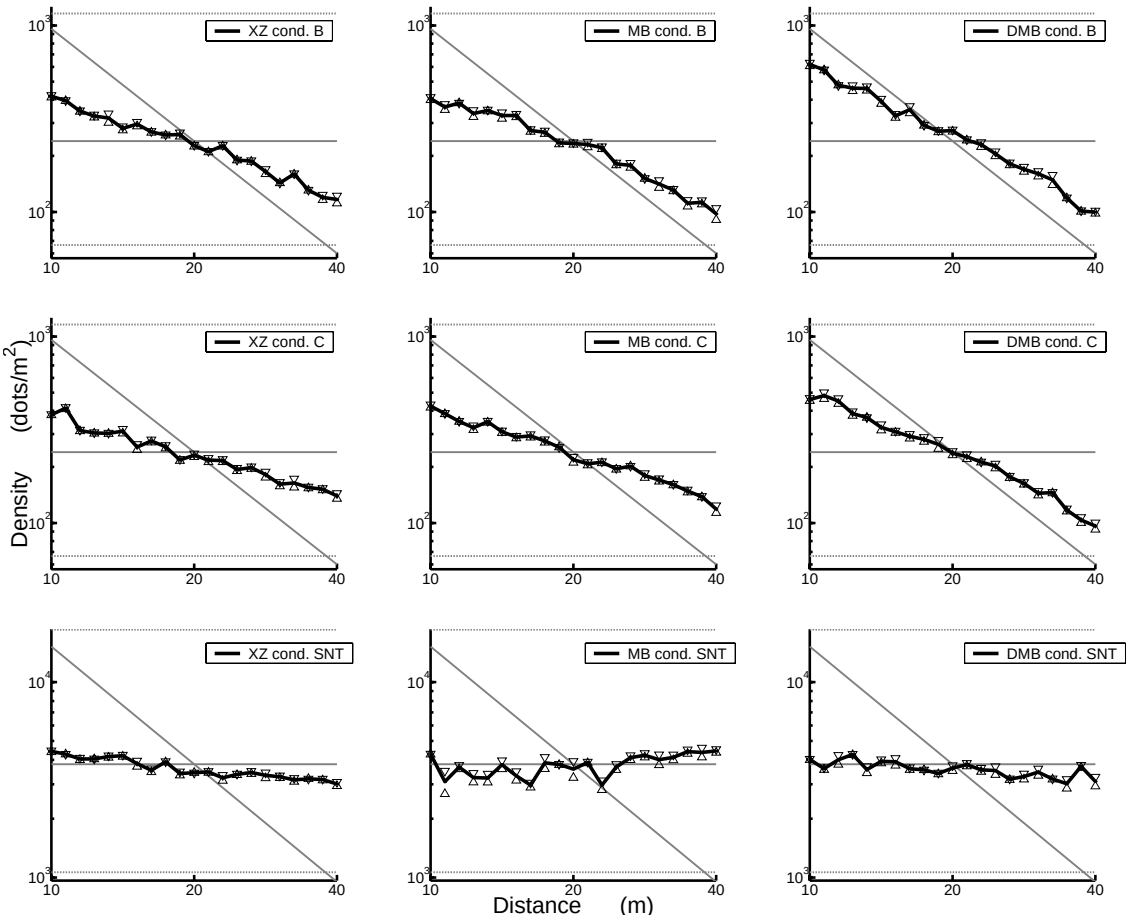

Fig. 14. Data collected for conditions B, C, and SNT of Experiment 2. Columns are data from the three observers, respectively. Rows show data for different conditions. The abscissa is distance of the test surface from the observer. The ordinate is the observer's setting of density for the test surface. Data points are from three independent trials; triangles mark standard errors. The dashed horizontal lines show the range of allowed responses. The solid horizontal line plots the density (in the world) of the reference surface, which is therefore the setting predicted by complete perceptual constancy. The solid diagonal line plots 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.

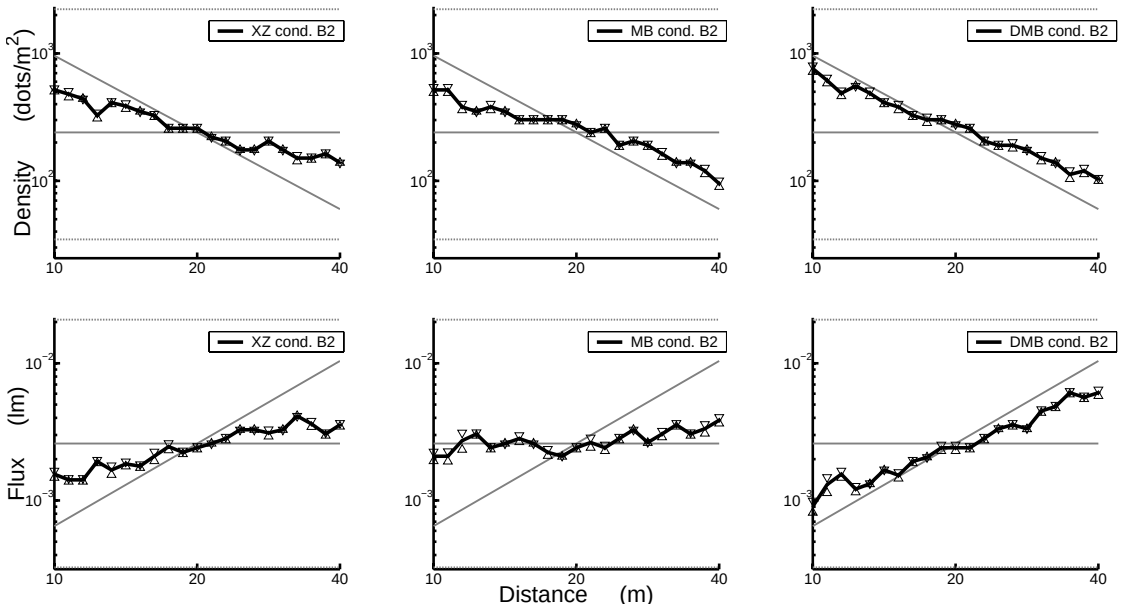

Fig. 15. Data collected for condition $B2$ of Experiment 2. Columns are data from the three observers, respectively. The top row shows the flux and the bottom one the density reported by the observers. Graph arrangement as in Figure 14.

the average of 3 independent measurements and the centers of the triangles indicate standard errors. The three flux settings were collected on the same trials as the three density settings, for a given observer at a given distance.

The data for condition $B3$ are shown in Figure 16. Each column represents the responses of one observer. The graphs are in the same format as Figure 14. In the figure, only the density settings are shown although that, in that condition, a keypress would increase or decrease both density and flux. The reason is that these increments or decrements were 100% correlated. In all graphs, each data point shows the average

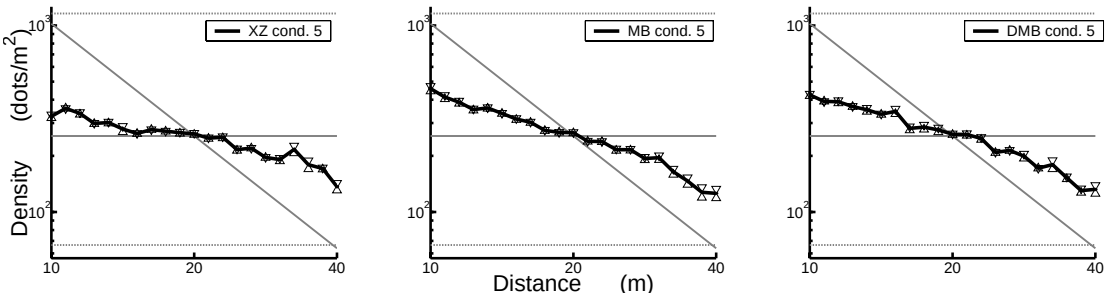

Fig. 16. Data collected for condition $B3$ of Experiment 2. Columns are data from the three observers, respectively. Only the density settings are shown because density and flux were 100% correlated in this condition. Graph arrangement as in Figure 14.

of 3 independent measurements and the 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-vs-distance lines in the log-log plots in Figures 14 to 16. 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 themselves have equal density.

Figure 17 plots these slopes for each of the three observers in each of the five 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 increasingly undercompensated for in conditions $B3$, C, B and $B2$, respectively.

The density settings for textures B and C (Figure 17) 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 is shown in the graphs by the fact that density in the world was not set to a constant value across changes in simulated distance: observers set near textures to be denser, and far textures to be sparser, than the texture on the reference surface. In terms of density in the *images*, simulated distance had a large effect, though not as large as predicted by simulated distance alone. This shortfall is not surprising because the perceptual consequences of manipulating simulated distance in computer displays are usually smaller than the manipulation predicts.[12]

Thus, one can interpret the data for textures B and C as showing that observers are able to accurately estimate the densities of textures in the world using an internal estimate of distance, but that in our experiments, this internal estimate varied less than the simulated distance.

Figure 17 also shows that observers' responses were different for SNT compared to their responses for textures B and C. Since the cutoff values for the realizable SNT were fixed, density and densintensity were both defined (they are proportional to one another) for the SNT test surface texture. The data show that settings for SNT, measured on the test surface in the world, varied little with changes in simu-

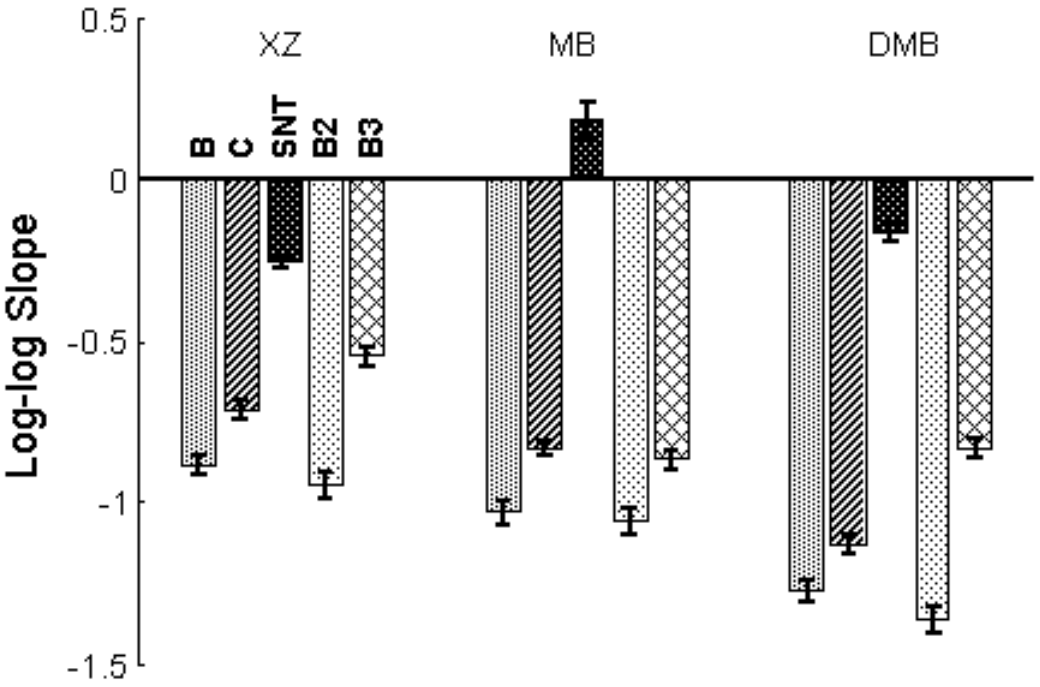

Fig. 17. Summary of data collected in Experiment 2. Each bar plots the slope of the regression line in the log-log plot of density setting vs simulated distance, from the data in Figures 14 to 16. Each group of five bars are data from one observer; the five bars show the slopes for conditions B, C, SNT, $B2$, and $B3$; the bars are colored from dark to light in that order. 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 are SEs based on total RMS error and the t-statistic

lated distance. The fact that density in the world was adjusted correctly for SNT, but not for textures B and C, is not an indication of better distance compensation for SNT, because images of SNT are invariant to changes in distance. Rather, we interpret these data as showing that the visual system correctly inferred that SNT should not look different at a new distance. Whatever the internal distance estimate, the transformation required to make the SNI of the test surface consistent with the SNI of the reference surface results in a test-surface image with the same statistics (densintensity) as the reference surface. To the extent that the visual system computes this transformation correctly, the observer would set densintensity to be the same in the image across changes in the distance of the test surface.

It is interesting to note that observers' density settings were not *completely* invariant to changes in simulated distance. This dependence was idiosyncratic: observer MB adjusted densintensity up with increasing distance, while by observers XZ and DMB adjusted densintensity down with increasing distance. One explanation is that, indeed, observers were not simply matching image statistics directly, but were instead estimating what the surface would look like at a different distance–and doing this slightly incorrectly.

The logic of this argument (namely that observers correctly make near-invariant settings for SNT across changes in distance) depends on our having successfully manipulated the internal estimate of distance in condition SNT. Is it possible that our simulated distance manipulation had a much greater effect on the internal distance estimate for textures B and C, than for SNT? If so, that could in itself explain why observers adjusted textures B and C to be different in the image of the test surface,

but not SNT. We used strong perspective cues because we wanted to ensure that surfaces would be seen as having different distances from the observer. Nevertheless, in conditions B and C there was an additional cue to distance change that was not present in SNT, namely the relative intensity of the IFS's in the images of the test and reference surfaces. If this dominated observers' density settings, then we might not be able to claim that internal distance estimates were manipulated successfully in condition SNT–and hence, the absence of a change in image statistics in that condition would be uninteresting.

We therefore ran two additional conditions using texture B, to measure the contribution of this cue to the density settings. In condition $B2$, observers adjusted both the density and the flux of the IFS's in the test surface. Intensity was not a distance cue in this condition; instead, it had to be adjusted according to the distance cue provided by perspective. In condition $B3$, flux was positively correlated with density, so flux and density were inversely related indicators of distance. In this condition, intensity became a "super cue": only when density was correct for the simulated distance was image intensity also correct, and only at that value was the image of the test surface consistent with the texture on the reference surface (i.e. for other settings, there was no distance of the test surface that could have given rise to the particular combination of density and intensity in the image[5]).

Observers' settings in condition $B2$ showed that they once again undercompen-

[5]This was also true for condition B, however, changing density by a given amount in that condition resulted in an image that deviated less from the nearest possible image than was the case in condition $B3$.

sated for simulated distance, by an amount similar to their undercompensation in condition B. Thus, intensity was not a major determinant of perceived distance in condition B. We conclude that internal distance estimates were probably similar in conditions SNT, B, and $B2$, so failure to appreciate simulated distance does not explain why the density of SNT was set correctly.

Density settings in condition $B3$ also showed undercompensation for simulated distance, but the settings were closer to prediction than for B and $B2$. Observers reported that this task seemed easier than conditions B and $B2$. In condition $B3$, there was additional information that could drive responses towards the predicted value, because there was only one density setting (namely, the setting consistent with simulated distance) for which the image was that of texture B at a new distance. All other settings represented a change in the texture itself, i.e. a change in the *ratio* of IFS flux to density. This ratio would normally be invariant in the image, across changes of distance. As a result, in condition $B3$ it was possible to set the density correctly for the simulated distance, without estimating the distance explicitly.

Condition $B2$ gives an indication of the extent to which observers can measure the ratio of intensity and density. In this condition, flux and density were adjusted separately, so the observer had an opportunity to make them both consistent with the same distance (i.e., to reproduce their ratio from the reference stimulus). The correlation between density and flux settings, across simulated distance, for observers XZ, DMB, and MB was -0.88, -0.62, and -0.82, respectively. Thus, simulated distance had a common effect on density and flux settings; this is also visible in the data of Figure 15 as anticorrelation in the residuals for the top and bottom graphs. The corre-

lation coefficients for the residuals in these graphs, across levels of simulated distance, were -0.68, -0.56, and -0.70 for the three observers, respectively. This correlation can be explained either by supposing that variance in the internal distance estimate had common direct effects on independent settings of intensity and density, or by supposing that the internal distance estimate was used to set one of these values, followed by setting the other value so that the ratio between intensity and density was the same as in the reference stimulus.

D. Experiment 3

In displays with real objects, perceived depth is consistent with binocular disparity cues.[13] In stereoscopic computer displays, however, perceived depth varies less than simulated depth.[12] 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,[14] 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,[15] the related cue of lack of change in monocular parallax (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 to other textures. While an SNI on a traditional display will still contain blur and accommodation cues, it would not contain density or intensity cues, and if small spots are used to simulate IFS's, an SNI would minimize texture element size and shape cues. Because physical images of SNT contain fewer cues that specify absence of depth, they ought, in principle, to cause the visual system to give greater weight to stereo depth cues, as compared with textures A, B and C. 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 textured with IFS's (texture A, B, C, or SNT). The patches were approximately circular in the images, and subtended $\approx 20^{\circ}$ of visual angle. Binocular disparity was the only cue that was manipulated explicitly. Vergence and vertical disparity were consistent with a viewing distance of $1m$, 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 IFS's 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 $\sqrt{}(HSR)$ and $1/\sqrt{}(HSR)$, respectively, where HSR was the desired horizontal size ratio between the left and right eyes' images, as determined by solving for HSR in the slant equation $S = -\tan^{-1}(1/(\mu \times \ln(HSR)))$, where μ was the vergence angle of the eyes at the $1m$ simulated viewing distance.[16] Thus the binocular disparity gradient was that of a flat surface with a slant of S degrees at $1m$. 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 *at* this mark during the experiment. Stimuli were prepared individually for each observer and were therefore correct for each observers' interocular distance. All 25×4 stimuli were presented once, in a random order, during a single session.

On the circular disk, densities and IFS fluxes for the four textures (A, B, C, SNT) were selected so that these textures would appear similar. For textures A and B, d_W was $240\, dots/m^2$ and flux $2.6 \times 10^{-3}\, lm$. For C, d_W was $240\, dots/m^2$ and flux values were from $5.2 \times 10^{-6}\, lm$ to $3.2 \times 10^{-3}\, lm$. For SNT, d_W was $3845\, dots/m^2$, $\alpha = 5.2 \times 10^{-6}$, and $\beta = 31.2 \times 10^{-3}\, lm$. As usual, for SNT the majority of the IFS's were too dim to be individually visible.

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

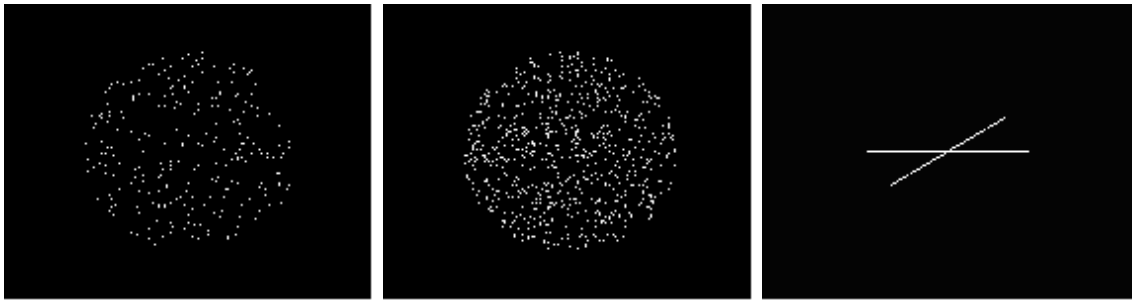

Fig. 18. Displays from Experiment 3. Left to right: (a) One image of a stereo pair that depicted a surface textured with texture B. (b) Same, for SNT. (c) The response stimulus used to indicate perceived slant. Observers adjusted the angle of the oblique line segment relative to the fixed horizontal segment, to indicate the slant they perceived in the stimulus.

response stimulus using keypresses. The response stimulus contained a pair of line segments,[17] and the angle between them was adjustable using keypresses. Observers indicated the perceived slant by adjusting this angle (see Figure 18). When the observer was content, he made a separate keypress and the computer recorded the slant setting and started the next trial.

3. Results

Figure 19 plots data collected from the two observers (the authors). In each graph, the four data series plot data for the four texture types, respectively. 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 slant specified by stereo cues.[17] These slopes are plotted (with error bars that show standard error) in

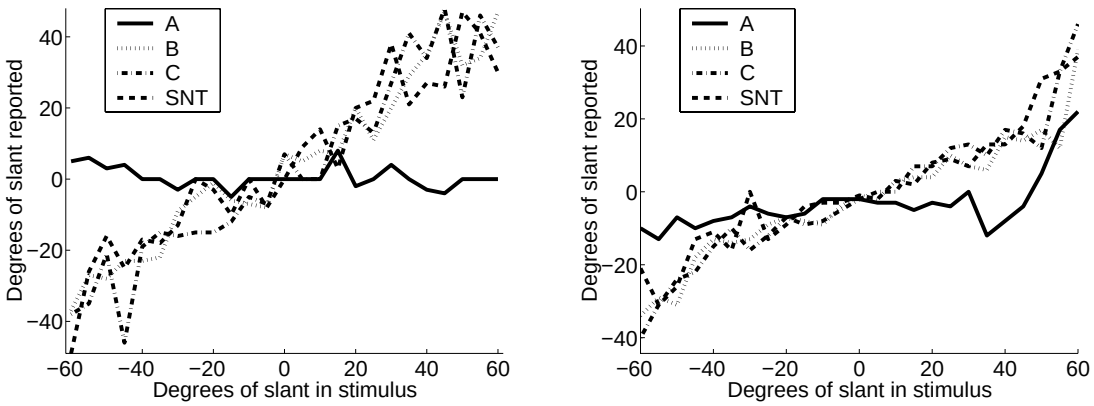

Fig. 19. Data from observers XZ (left) and BTB (right) in Experiment 3. The abscissa plots the slant in the stimulus that was specified by disparity cues. The ordinate plots the observer's slant response. Different linestyles correspond to different textures, as shown in the legend.

the bargraph shown by Figure 20.

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 from stimulus to percept, and to what extent the mapping from percept to response.[18] However, the data in

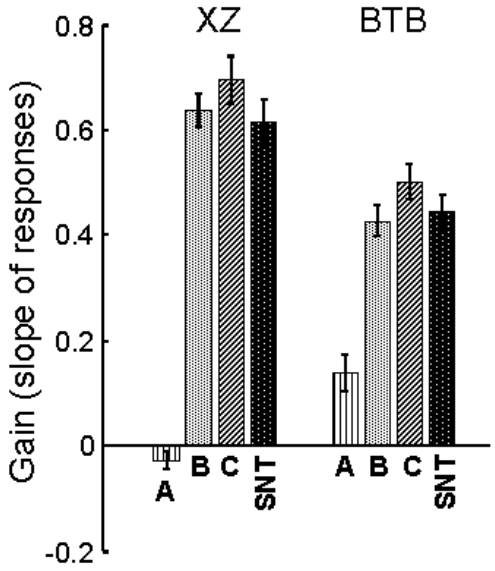

Fig. 20. Gain of the slant response, relative to slant specified by disparity. Data from the two observers, for the four conditions. Each bar is the coefficient of regression for one of the data series in Figure 19. Error bars are standard errors.

Figure 20 makes it clear that texture A evoked very little perceived slant, and that textures B, C, and SNT evoked significant perceived slant, the magnitude of which depended on the disparity gradient. Textures B, C, and SNT evoked slant percepts that were not measurably different from one another.

Why was perceived slant so small for texture A? The visual system presumably knows that a square lattice of dots in the image is most likely to have been caused by an unslanted, square lattice of dots in the world. The alternative scenes that might have produced this image all require contrived surface patterns, on surfaces that, by coincidence, have just the right depth profile to cancel out the pattern contrivance in the image. The prior probability for slant is highest for unslanted objects (optical slant in the world is distributed as $cos(slant)$ if surfaces have random orientations in space), homogeneously textured surfaces are more common than surfaces patterned with texture gradients, and coincidental viewpoints are unlikely. These three factors conspire with accommodation and residual texture element size and shape cues, to prevent the stereo signals from prevailing during construction of the percept. In Bayesian terms, the likelihoods and priors for stimulus A both contribute to a posterior probability that strongly favors the unslanted, regular surface.[19]

We predicted that perceived slant would be greater for SNT than for textures B and C. This was not the case. To understand why this might be the case, let us consider the sources of information available in textures B and SNT. Binocular disparity specified the same slants for both textures. The outlines of the stimuli were always circular in the image, so it specified a slant of zero for both textures, under the reasonable assumption that circles are the most commonly observed ellipses.

Accommodative demand was approximately constant across the display, so it also specified a slant of zero for both stimuli. Very small dots were used, but to the extent that a gradient in dot size was visible across the display, it specified a slant of zero. The texture gradient was defined for texture B but not for texture SNT. Our hypothesis was that the visual system should interpret the texture gradient of zero as an additional cue to slant for texture B, that was not present for SNT, resulting in perceived slants of smaller magnitude for texture B.

For texture A, it is reasonable to suppose that texture was a strong indicator that the surface had zero slant. To what extent was this the case for texture B? Would a truly slanted surface give rise to images with measurably different image statistics? If not, there would be no reason to suppose that a texture gradient of zero would be given weight by the visual system. The stimuli in this experiment subtended 18° of visual angle. An object slanted by 60°, seen within this window, is 1.8 times farther on the right side than the left, and is slanted at 60°, ±9° at the two sides of the window, respectively, resulting in a density ratio between the extreme sides of the display of $1.7^2 \times \cos(51°)/\cos(69°) = 5.7$. The intensity ratio between sides of the window would, correspondingly, be 1/5.7. Variations of this magnitude would be easily visible in our displays. Thus, density and intensity cues should have been reliable indicators that the slant of texture B was 0°, rather than 60∘, when stereo specified 60°. In that case, texture cues in the image of texture B should have been given significant weight, so texture B should have appeared less slanted than SNT, but this was not the case.

However, three additional factors must be considered. First, the internal representation of distance, that was used to scale disparity, may have been less than the

$1\,m$ specified by vergence and vertical disparity. If it were $50, cm$ instead, then the maximum disparity gradient in the display ($HSR = 1.113$ for PD $= 6.2\,cm$) should be interpreted not as 60°, but rather as 41°. At this slant, the density ratio between extreme sides of the display would be only 2.6. More generally, texture cues provide less information about changes in slant for surfaces with small slants than for surfaces with large slants.[19] It is possible that the magnitude of the texture gradient itself has a disproportionately large impact on the extent to which the visual system relies on texture as a slant cue. In that case, one might predict the following in situations where texture provides weak information about slant: because the texture gradient is zero, it is given little weight. If so, then texture would be given little weight for both textures B and SNT.

Second, we saw in Experiment 2 that observers did not keep SNT completely invariant in the image, across changes in distance. Thus it would be reasonable to suppose the visual system doesn't correctly compute expected changes in the images for SNT across changes in slant, either. Although changes in slant do not change the image statistics for SNT, the visual system might not know this; it might erroneously behave as though texture were a cue to slant for SNT, that indicated zero slant.

Finally, it is possible that slant was underestimated because slant cues provided by the CRT screen had a large effect. These cues were the same for SNT and texture B. If they played a large role in determining perceived slant, then the relative weight accorded to the density gradient in texture B would have been correspondingly reduced, and this could make it difficult to see the effect of presence (texture B) vs. absence (SNT) of a texture density gradient. Observers are able to use this cue to

make monocular judgments of the slant of a CRT screen[16] and the same pattern of retinal disparities evokes greater perceived slant when the CRT screen agrees with the stereoscopically specified slant, as compared to when the screen is unslanted.[20,21]

The foregoing analyses suggest that other experiments might be able to show greater stereo depth in SNT than texture B. For example, if observers were asked to estimate depth intervals instead of slant, in displays where disparity specified a step change in distance from one part of the display to another, then texture might have been a strong cue to the absence of depth for texture B, strong enough to result in greater perceived depth from stereo for SNT than for texture B.

Finally, we allowed unlimited views of displays. Imposing a time limit might, in principle, reveal a difference between SNT and textures B and C; subjectively, observer XZ felt that slant was more readily visible for SNT, and that perceived slant grew over time more for textures B and C than texture SNT. So it is possible that the initial estimate of slant was dominated by texture, with additional weight being given to stereo over time. In that case, it is conceivable that small differences between perceived slant in textures B or C, and SNT, might have been resolvable using shorter display durations.

4. Summary

The SNT is a novel texture, that does not convey information about surface distance or slant when imaged. SNT is not realizable in practice. To implement it, compromises were necessary: the distribution of IFS fluxes must be truncated at both ends, and finite-size display elements must be used to represent the IFS's.

Experiment 1 showed that the ability of SNT to hide changes in distance survived the truncation of the distribution of IFS fluxes. Experiment 2 showed that observers have significant ability to match textures in the world across changes in perceived distance, and that they appreciated, at least to some extent, the fact that SNI's generated by a particular 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 isolating binocular disparity (or motion parallax) depth cues remains to be seen. In principle, SNT allows the experimenter to create stimuli that do not contain pictorial cues to depth. We could not demonstrate this in Experiment 3, because unavoidable cues in computer displays (such as texture element size and accommodation) limited the effectiveness of disparity as a cue to depth, and/or because observers interpreted SNI as indicating zero slant when in fact it was uninformative. In Experiment 2, observers set SNI to be nearly, but not exactly equal to one another across changes in distance. It is possible that through training, during which observers interact with real or simulated objects that are painted with SNT, the visual system might learn to better appreciate that image statistics for SNT are invariant to changes in distance and slant.

Acknowledgments

X. Zabulis was supported by the National Science Foundation as a fellow at the Institute for Research in Cognitive Science at the University of Pennsylvania. The authors thank Prof. Martin S. Banks for helpful discussions with the 2nd author during con-

ception of the problem and its possible solutions, David Brainard for comments on the properties of SNT, Benjamin Rosenau for his help constructing the metal foil demonstration of SNT, and Daniel Matza-Brown and Michael Brunswick for serving as observers.

References

1. B. Julesz. Binocular depth perception without familiarity cues. *Science*, 145:356–62, 1964.
2. J.J. Gibson. *The Perception of the Visual World*. Houghton-Mifflin, Boston, 1950.
3. J.J. Gibson. *The ecological approach to visual perception*. Houghton Mifflin, Boston, 1979.
4. J. Malik and R. Rosenholtz. Computing local surface orientation and shape from texture for curved surface. *International Journal of Computer Vision*, 23(2):149–168, 1997.
5. R. Rosenholtz and J. Malik. Surface orientation from texture : Isotropy or homogeneity (or both)? *Vision Research*, 37(16):2283–93, 1997.
6. B.A. Dosher, G. Sperling, and S. Wurst. Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research*, 26(6):973–90, 1986.
7. V. Nalwa. *A Guided Tour of Computer Vision*. Addison-Wesley, 1993.
8. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2000.
9. H. W. Olbers. Über die durchsichtigkeit des weltraums. *Astronomisches Jahrbuch*

für das Jahr 1826, 1823. Engl. transl. 1826 On the transparency of space Edinburgh New Phil. J. 1.

10. D. Knill and W. Richards, editors. *Perception as Bayesian Inference*. Cambridge University Press, Cambridge, 1996.
11. D. H. Brainard. The Psychophysics Toolbox. *Spatial Vision*, 10:433–436, 1997.
12. J. M. Loomis and J. M. Knapp. *Virtual and Adaptive Environments*, chapter Visual perception of egocentric distance in real and virtual environments. Erlbaum, Mahwah NJ, (in press).
13. J. P. Frisby, D. Buckley, and P. A. Duke. Evidence for good recovery of lengths of real objects seen with natural stereo viewing. *Perception*, 25:129–54, 1996.
14. W. C. Gogel. Equidistance tendency and its consequences. *Psychological Bulletin*, 64:153–163, 1965.
15. G. Mather. The use of image blur as a depth cue. *Perception*, 26:1147–1158, 1997.
16. B.T. Backus, M.S. Banks, R. van Ee, , and J.A. Crowell. Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, 39(6):1143–70, 1999.
17. R. van Ee and C. J. Erkelens. Temporal aspects of binocular slant perception. *Vision Research*, 36(1):43–51, 1996.
18. M. S. Banks and J. L. Dannemiller. *Visual psychophysics. Handbook of Infant Perception*, pages 115–184. Academic Press, New York, 1987.
19. D. C. Knill. Surface orientation from texture: Ideal observers, generic observers

and the information content of texture cues. *Vision Research*, 38(11):1655–1682, 1998.

20. S. J. Watt, M. S. Banks, M. O. Ernst, and J. M. Zumer. Screen cues to flatness do affect 3d percepts. *Journal of Vision*, 2(7):297a, 2002. Conference Abstract: VisionSciences Annual Meeting.

21. S.J. Watt, K. Akeley, and M.S. Banks. Focus cues to display distance affect perceived depth from disparity. page TA66, 2003. Conference Abstract. To also appear in the Journal of Vision.