Discrimination of textures with spatial correlations and multiple gray levels

JONATHAN D. VICTOR,[1,*] SYED M. RIZVI,[1,2] JACOB W. BUSH,[1,3] AND MARY M. CONTE[1]

[1]*Feil Family Brain and Mind Research Institute, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10065, USA*
[2]*Currently with Centerlight Healthcare, 136-65 37th Ave., Flushing, NY 11354, USA*
[3]*Currently with Shopify, 151 O'Connor St Ground floor, Ottawa, ON K2P 2L8, Canada*
**jdvicto@med.cornell.edu*

Abstract: Analysis of visual texture is important for many key steps in early vision. We study visual sensitivity to image statistics in three families of textures that include multiple gray levels and correlations in two spatial dimensions. Sensitivities to positive and negative correlations are approximately independent of correlation sign, and signals from different kinds of correlations combine quadratically. We build a computational model, fully constrained by prior studies of sensitivity to uncorrelated textures and black-and-white textures with spatial correlations. The model accounts for many features of the new data, including sign-independence, quadratic combination, and the dependence on gray level distribution.

1. Introduction

One of the strategies that the visual system uses to grapple with the complexity of analyzing natural sensory signals is to organize this analysis according to groups of attributes – for example, orientation, color, motion, and depth [2]. For these classical submodalities of spatial vision, this organizational strategy has well-recognized anatomical underpinnings, both at the level of specialization of cortical areas and the tuning properties of their component neurons [2-6].

Although the specialization of visual areas and the independence of processing within submodalities is far from absolute [7-10], it is clear that computational "factoring" is an important principle. That is, while a neuron may be tuned to more than one submodality of spatial vision (e.g., its response may depend both on color and orientation), its selectivity can often be understood by considering one submodality at a time. Conversely, it is rare to find a neuron whose preferred spatial orientation changes as a function of the chromaticity of the grating used to probe it. Intuitively, this arrangement is a natural consequence of parallel visual streams and simplifies the logic needed to read out a pattern of neural activity.

Here, we ask whether this computational principle generalizes to another aspect of spatial vision – visual texture. There are two ways in which texture differs from the classic submodalities, and thus, two reasons that this generalization is not a foregone conclusion. First, the connection between perceptual sensitivities and tuning properties of individual neurons is likely to be less direct than for the classic submodalities: texture, by its nature, cannot be signaled by a small number of localized receptive fields, as it is a statistical characterization of an image across an extended region.

Second, the domain of visual texture is high-dimensional. The reason for this is that any statistic that measures the joint probability of a set of luminance values in any given spatial configuration is, potentially, a perceptual dimension for texture, i.e., a parameter for which visual sensitivity may be tuned. Within this vast range of possibilities, visual sensitivity is highly selective – but nevertheless, there are a large number of such image statistics for which visual sensitivity is substantial. [11-14].

To address whether computational "factoring" extends to texture, we measure visual sensitivity to image statistics that incorporate two aspects of texture that are typically studied separately: the distribution of luminance levels, and the spatial organization of the correlations. We then construct a model for these sensitivities. The model has the familiar "back-pocket" structure [15], but each channel of the model posits a specific way in which analysis of image statistics can be separated into a component that is sensitive to the distribution of luminance levels, and a component that is sensitive to spatial configuration. The model's parameters are then constrained by requiring it to account for two complementary, pre-existing psychophysical datasets that do not overlap the current study: sensitivity to differences in the luminance histogram in textures with no spatial structure [14, 16, 17], and sensitivity to differences in spatial configuration in textures with only two luminance levels [13, 18]. Because of the model's simple structure, it can be fully constrained by this requirement, with no free parameters. We find that the model provides an approximate account of the new psychophysical measurements, in terms of relative sensitivities to different kinds of image statistics and how different image statistics combine.

2. Materials and methods

Our overall experimental strategy is to use synthetic visual textures to measure visual sensitivity to image statistics and their combinations. As in previous work, a texture is formally defined as an ensemble of infinitely large images, with the requirement that its statistics can be equivalently estimated either by averaging a single sample over all of space, or averaging across many examples of a finite patch[18, 19]; our stimuli consist of random samples drawn from such an ensemble. The textures we consider here are all composed of monochrome checks, and the statistics we consider are all local correlations, i.e., the average value, across the ensemble, of a product of luminances of checks at specific relative displacements.

Despite these restrictions, a practical challenge remains: the number of image statistics required to specify a texture is enormous [20, 21]. This challenge, along with a range of theoretical considerations [18, 22-25], motivates the adoption of the "maximum-entropy" approach used here: a small number of image statistics are specified explicitly, and the texture ensemble is constructed to be as random as possible, given these constraints.

In this work, the constraints are the luminance distribution and correlations of checks within a 2×2 neighborhood. We use a 2×2 region (here, and in previous studies that this work builds on [13, 18, 26-28]) because it is the smallest region that enables specification of textures with contours and corners in multiple directions, as well as T-junctions and X-junctions.

This approach provides a practical dimension reduction and also one which, perhaps surprisingly, is related to the statistics of natural images[24, 25]. Our approach is related to, but distinct from, the "FRAME" approach to texture synthesis of Zhu et al. [22]. While both are maximum-entropy approaches, FRAME uses constraints that are neurally-inspired linear spatial filters applied to the image (and thus, also encompasses the original texton approach of Julesz [29, 30]); here, the constraints are nonlinear combinations of local luminances.

Psychophysical measurements of sensitivity to individual image statistics and their combinations were made by using the texture segmentation task introduced by Chubb et al. [17] and used in many previous studies in our lab [13, 18, 26-28] for black-and-white textures. Here we describe an extension of this approach to multiple gray levels. We then detail the psychophysical task, experimental procedure, and data analysis. Construction of the textures is detailed in the Supplemental Document ("Specification and construction of textures"). This construction maintains the maximum-entropy property of the black-and-white construction [18]: textures are as random as possible for the image statistics that are specified. Because of this maximum-entropy property, the textures contain the minimal visual structure that is required to achieve the specified image statistics.

A portion of the psychophysical data presented in Experiments 1 and 2 has also been presented in [23], but without many of the experimental details.

2.1 Stimuli for experiments 1 and 2

Experiments 1 and 2 extend the analysis of black-and-white textures [18] to textures with three luminance levels. In the binary context, we developed a coordinate system for image statistics that comprehensively described all kinds of correlations within a 2×2 neighborhood of checks; we now expand the coordinate system to take into account multiple luminance levels.

In the case of black-and-white textures, image statistics are grouped according to "order", i.e., the number of checks that are multiplied to calculate the statistic. For example, the first-order statistic specifies the luminance distribution of individual checks, and the second-order statistics describe the pairwise correlation of luminances in a pair of checks. There are four second-order statistics, since there are four kinds of two-check correlations to be considered: between two checks that are adjacent horizontally, vertically, and along each of the two diagonals. Each statistic thus specifies the expected value of the product of the luminances of horizontally adjacent, vertically adjacent, or diagonally adjacent check pairs, averaged across all samples of the texture. Similarly, there are four third-order statistics, corresponding to the four ways of selecting three checks within a 2×2 neighborhood; each statistic specifies the expected value of the product of three luminances. Finally, there is one fourth-order statistic; it specifies the correlation among all four checks, i.e., the product of the four luminances.

To extend this scheme to multiple luminance levels, we group image statistics according to order (the number of checks whose luminances are multiplied), and subdivide each order according to the spatial configuration of the checks. However, each of these subdivisions now becomes a family of statistics, as more than one parameter is needed to describe the correlations among checks in a given configuration (Table 1). Furthermore, each family (other than first-order) subdivides into independent genera based on the rule that links the luminance values within each configuration; each genus in turn contains several species -- the individual statistics that specify the distribution of values created by this rule. (We are borrowing the standard taxonomic nomenclature – order, family, genus, and species -- for a hierarchy that arises out of mathematical considerations, but we do not intend to imply a hierarchical structure for the visual computations). The need for this structure becomes apparent when we consider three or more luminance levels and statistics of order two or more; these play a key role in the first two experiments.

First-order statistics

First-order statistics describe the distribution of luminance level values assigned to each check. When three levels are present, the distribution is specified by the probability that a check is black (0), gray (1), or white (2). This is a three-element vector, $(p(0), p(1), p(2))$, which we denote as $\vec{\sigma}_{(1)}$. Since the probabilities of black, gray, and white checks must sum to 1, there are two degrees of freedom, so this family requires two image statistics – the "species" within this family.

We represent these two degrees of freedom as barycentric coordinates (page 216 of [1]) of a triangular domain (Fig. 1), whose vertices correspond to textures that are all black $\vec{\sigma}_{(1)} = (1,0,0)$, all gray $\vec{\sigma}_{(1)} = (0,1,0)$, or all white $\vec{\sigma}_{(1)} = (0,0,1)$. The centroid of the triangle, $\vec{\sigma}_{(1)} = (1/3, 1/3, 1/3)$, corresponds to a texture in which each gray level occurs 1/3 of the time, and there are no spatial correlations. Note that in the black-and-white case,

Table 1. Texture Coordinates

order	block probabilities	Fourier coordinates	reduced Fourier coordinates	barycentric coordinates	binary coordinates
1	$p\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$	$\varphi\begin{pmatrix} s_1 & s_2 \\ s_3 & s_4 \end{pmatrix}$	$\varphi(s_1)$	$\vec{\sigma}_{(1)}$	$\vec{\sigma}_{(1)} = \left(\frac{1-\gamma}{2}, \frac{1+\gamma}{2}\right)$
2			$\varphi\begin{pmatrix} s_1 & s_2 \end{pmatrix}$	$\vec{\sigma}_{(1 \ \ s_2)}$	$\vec{\sigma}_{(1 \ \ 1)} = \left(\frac{1+\beta_{-}}{2}, \frac{1-\beta_{-}}{2}\right)$
			$\varphi\begin{pmatrix} s_1 \\ s_3 \end{pmatrix}$	$\vec{\sigma}_{\begin{pmatrix} 1 \\ s_3 \end{pmatrix}}$	$\vec{\sigma}_{\begin{pmatrix} 1 \\ 1 \end{pmatrix}} = \left(\frac{1+\beta_{\vert}}{2}, \frac{1-\beta_{\vert}}{2}\right)$
			$\varphi\begin{pmatrix} s_1 & \\ & s_4 \end{pmatrix}$	$\vec{\sigma}_{\begin{pmatrix} 1 & \\ & s_4 \end{pmatrix}}$	$\vec{\sigma}_{\begin{pmatrix} 1 & \\ & 1 \end{pmatrix}} = \left(\frac{1+\beta_{\backslash}}{2}, \frac{1-\beta_{\backslash}}{2}\right)$
			$\varphi\begin{pmatrix} & s_2 \\ s_3 & \end{pmatrix}$	$\vec{\sigma}_{\begin{pmatrix} & 1 \\ s_3 & \end{pmatrix}}$	$\vec{\sigma}_{\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}} = \left(\frac{1+\beta_{/}}{2}, \frac{1-\beta_{/}}{2}\right)$
3			$\varphi\begin{pmatrix} s_1 & s_2 \\ s_3 & \end{pmatrix}$	$\vec{\sigma}_{\begin{pmatrix} 1 & s_2 \\ s_3 & \end{pmatrix}}$	$\vec{\sigma}_{\begin{pmatrix} 1 & 1 \\ 1 & \end{pmatrix}} = \left(\frac{1-\theta_{\ulcorner}}{2}, \frac{1+\theta_{\ulcorner}}{2}\right)$
			$\varphi\begin{pmatrix} s_1 & s_2 \\ & s_4 \end{pmatrix}$	$\vec{\sigma}_{\begin{pmatrix} 1 & s_2 \\ & s_4 \end{pmatrix}}$	$\vec{\sigma}_{\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}} = \left(\frac{1-\theta_{\urcorner}}{2}, \frac{1+\theta_{\urcorner}}{2}\right)$
			$\varphi\begin{pmatrix} s_1 & \\ s_3 & s_4 \end{pmatrix}$	$\vec{\sigma}_{\begin{pmatrix} 1 & \\ s_3 & s_4 \end{pmatrix}}$	$\vec{\sigma}_{\begin{pmatrix} 1 & \\ 1 & 1 \end{pmatrix}} = \left(\frac{1-\theta_{\llcorner}}{2}, \frac{1+\theta_{\llcorner}}{2}\right)$
			$\varphi\begin{pmatrix} & s_2 \\ s_3 & s_4 \end{pmatrix}$	$\vec{\sigma}_{\begin{pmatrix} & 1 \\ s_3 & s_4 \end{pmatrix}}$	$\vec{\sigma}_{\begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}} = \left(\frac{1-\theta_{\lrcorner}}{2}, \frac{1+\theta_{\lrcorner}}{2}\right)$
4			$\varphi\begin{pmatrix} s_1 & s_2 \\ s_3 & s_4 \end{pmatrix}$	$\vec{\sigma}_{\begin{pmatrix} 1 & s_2 \\ s_3 & s_4 \end{pmatrix}}$	$\vec{\sigma}_{\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}} = \left(\frac{1+\alpha}{2}, \frac{1-\alpha}{2}\right)$
parameter count	G^4 real	G^4 complex	$G(G-1)\times$ (G^2+G-1) complex	$G\times$ (G^2+G-1) barycentric vectors	10
random texture	all $\frac{1}{G^4}$	all 0 except $\varphi\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 1$	all 0	all $\frac{1}{G}$	all 0

Parameterization of local image statistics in terms of block probabilities. G (columns 2 through and 5) is the number of gray levels. For $G = 3$, the barycentric coordinates correspond to triangular domains, as shown in Figs. 1-3. For $G = 2$, these domains are one-dimensional, and correspond to the image statistics of [18], as shown in the rightmost column. The rows of the last three columns correspond to families of statistics.

there was only one degree of freedom for first-order statistics– since the fraction of black and white checks must sum to 1. This single degree of freedom was captured by a single parameter γ, where $(1+\gamma)/2$ is the probability of white checks, and $(1-\gamma)/2$ is the probability of black checks. The final two columns of Table 1 specify the correspondence between the barycentric coordinates, which apply to any number of gray levels, and the binary coordinates introduced in [18] and used in previous studies.

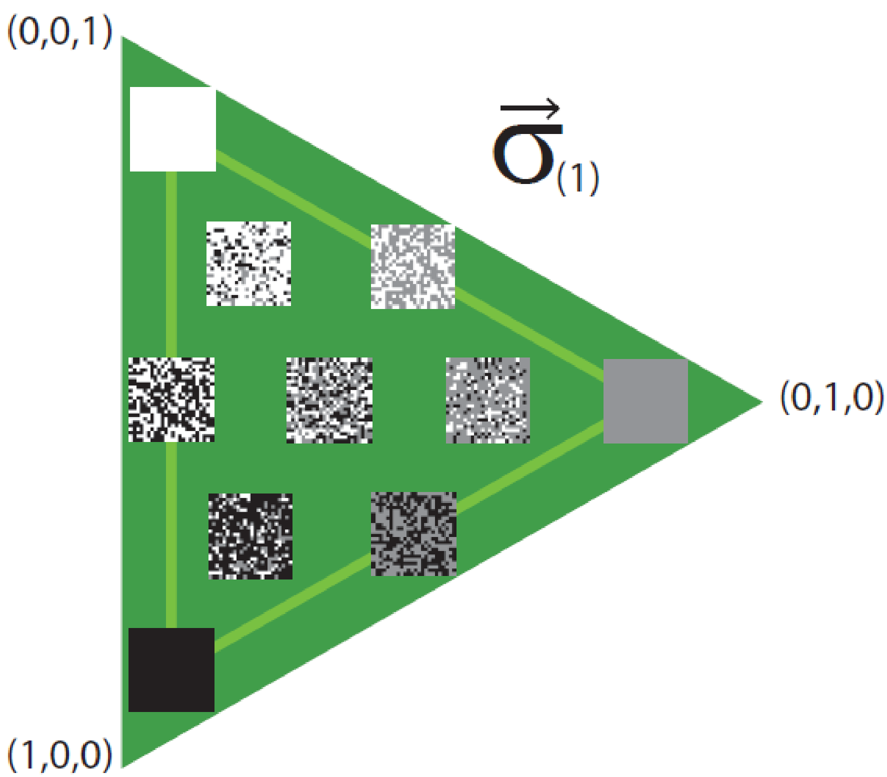

Fig. 1. The domain of the first-order statistic $\vec{\sigma}_{(1)}$ for three-level textures. $\vec{\sigma}_{(1)}$ is a three-element vector whose entries correspond to the probability of black, gray, and white checks, respectively. Since these three values must sum to 1, they can be considered as barycentric coordinates [1] (page 216) for a triangle. The vertices of the triangle are the extreme points of the domain, and correspond to the probability distributions that are all black $\vec{\sigma}_{(1)} = (1,0,0)$, all gray $\vec{\sigma}_{(1)} = (0,1,0)$, or all white $\vec{\sigma}_{(1)} = (0,0,1)$. The centroid of the triangle, which corresponds to $\vec{\sigma}_{(1)} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, corresponds to a random texture.

Second-order statistics

For second-order statistics, we detail the family of statistics that describe correlations between two horizontally-adjacent checks; the other three families of second-order statistics, which describe correlations in vertical and diagonal directions, are handled similarly.

There are nine ways that a pair of horizontally-adjacent checks can be colored by three

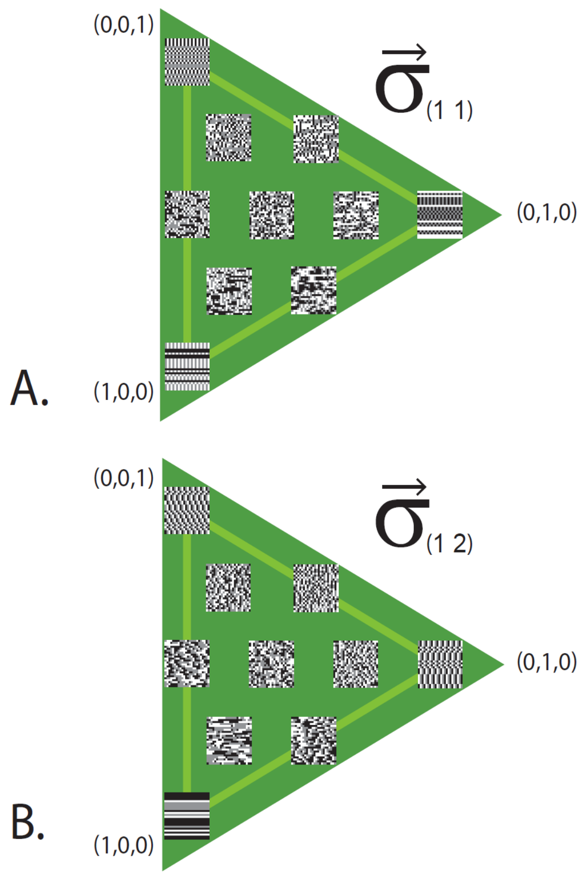

Fig. 2. The domains of the second-order statistics $\vec{\sigma}_{(1\ \ 1)}$ (A) and $\vec{\sigma}_{(1\ \ 2)}$ (B) that capture the pairwise correlation of luminance levels in horizontally-adjacent checks. Within each domain, a three-element vector $\vec{\sigma}_{(1\ \ s)}$ ($s=1$ in panel A, $s=2$ in panel B) describes the kind of horizontal correlation. Specifically, the elements of $\vec{\sigma}_{(1\ \ s)}$ give the probability distribution of $A_1 + sA_2 \pmod 3$, where A_1 and A_2 are the luminance values of the checks (0 for black, 1 for gray, 2 for white). Since the three values of each $\vec{\sigma}_{(1\ \ s)}$ are a probability distribution and therefore sum to 1, the domain of each vector forms a triangle (as in Fig. 1). The vertices of the triangle, $\vec{\sigma}_{(1\ \ s)} = (1,0,0)$, $\vec{\sigma}_{(1\ \ s)} = (0,1,0)$, and $\vec{\sigma}_{(1\ \ s)} = (0,0,1)$, correspond to textures in $A_1 + sA_2$ is either always 0, always 1, or always 2. Therefore, the textures at the vertices have rows that are completely determined by their initial check. Also as in Fig 1. (and as in all other triangular domains), the centroid of the triangle corresponds to a random texture.

luminance levels. We denote these nine probabilities by $p\left(A_1 \quad A_2\right)$, where A_1 and A_2 denote the luminances (0,1, or 2) assigned to the two checks. These nine probabilities must sum to 1. There are additional constraints implied by the first-order statistics. For example, summing $p\left(A_1 \quad A_2\right)$ over A_2 must yield $p\left(A_1\right)$, and summing $p\left(A_1 \quad A_2\right)$ over A_1 must yield $p\left(A_2\right)$. Consequently (see Supplemental Document), there are four degrees of freedom for the second-order statistics that describe horizontal correlations.

These four parameters can be grouped into two independent triangular domains (Fig. 2), the "genera" for this family. The first domain (Fig. 2A) links luminance values by constraining the distribution of $A_1 + A_2 \pmod 3$ (here, as is standard, "mod n" denotes the remainder after division by n); the second (Fig. 2B) links luminance values by constraining the distribution of $A_1 + 2A_2 \pmod 3$. In each case, the possible values of the sum are 0, 1, or 2, so the distribution of the sum is described by a three-element vector of elements that sum to 1. We denote these vectors as $\vec{\sigma}_{(1\ \ 1)}$ for $A_1 + A_2$ and $\vec{\sigma}_{(1\ \ 2)}$ for $A_1 + 2A_2$: the subscripts indicate the values of the multipliers and their positions within the 2×2 neighborhood. As for the first-order statistic $\vec{\sigma}_{(1)}$, the vertices of each triangle correspond to extremes of the distribution, in which only one value of the sum occurs. The centroid of the triangle corresponds to the random texture, where each value of the sum has probability $1/3$.

Inspection of the texture samples at the vertices of these triangular domains shows that $\vec{\sigma}_{(1\ \ 1)}$ and $\vec{\sigma}_{(1\ \ 2)}$ describe quite different aspects of pairwise correlations. For $\vec{\sigma}_{(1\ \ 1)}$ (Fig. 2A), each extreme texture consists of two kinds of rows: rows that contain only one luminance level, and rows that contain alternation of the other two levels. For example, for $\vec{\sigma}_{(1\ \ 1)} = (1,0,0)$ (bottom vertex of the triangle in Fig. 2A), luminance values of horizontally-adjacent check pairs must sum to 0 (mod 3). Thus, the only allowed pairs are $(0,0)$, $(1,2)$, and $(2,1)$, so every row is either only black, or alternating white and gray. Similarly, for $\vec{\sigma}_{(1\ \ 1)} = (0,0,1)$ (top vertex of the triangle in Fig. 2A), luminance values of horizontally-adjacent checks must sum to 2 (mod 3). Thus, the allowed pairs are $(1,1)$, $(2,0)$, and $(0,2)$ and every row is either only gray, or alternating white and black.

In contrast, the textures for $\vec{\sigma}_{(1\ \ 2)}$ have very different characteristics (Fig. 2B). At the bottom vertex, $\vec{\sigma}_{(1\ \ 2)} = (1,0,0)$ specifies that $A_1 + 2A_2$ is always equal to 0 (mod 3). This is equivalent to $A_1 = A_2 \pmod 3$, so all rows contain just one luminance level. The other two vertices of the domain correspond to rows that cycle between the colors. For the right vertex, $\vec{\sigma}_{(1\ \ 2)} = (0,1,0)$, the coloring (reading from left to right) cycles from white to gray to black, since $\vec{\sigma}_{(1\ \ 2)} = (0,1,0)$ means that $A_1 + 2A_2 = 1 \pmod 3$, so $A_2 = A_1 - 1 \pmod 3$ and the allowed pairs are $(2,1)$, $(1,0)$, and $(0,2)$. For the top vertex, $\vec{\sigma}_{(1\ \ 2)} = (0,0,1)$, the coloring cycles in the opposite order, since $\vec{\sigma}_{(1\ \ 2)} = (0,0,1)$ means that $A_1 + 2A_2 = 2 \pmod 3$, so $A_1 = A_2 + 2 = A_2 - 1 \pmod 3$, yielding the allowed pairs $(0,1)$, $(1,2)$, and $(2,0)$.

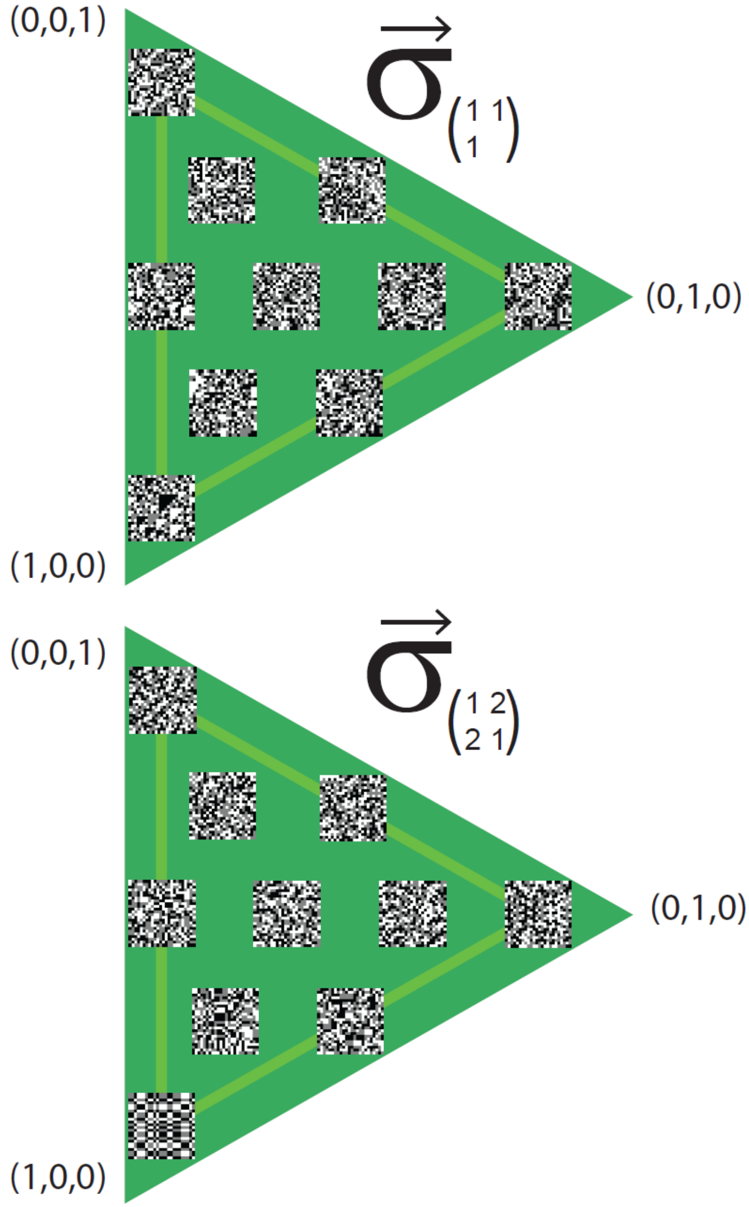

Fig. 3. A. The domain of the third-order statistics $\vec{\sigma}_{\begin{pmatrix}1&1\\1&\end{pmatrix}}$ that describes the correlation among the three checks $\begin{pmatrix}A_1&A_2\\A_3&\end{pmatrix}$, according to the distribution of $A_1+A_2+A_3 \pmod 3$. B: The domain of the fourth-order statistics $\vec{\sigma}_{\begin{pmatrix}1&2\\2&1\end{pmatrix}}$ that describe correlations of luminance levels among the four checks $\begin{pmatrix}A_1&A_2\\A_3&A_4\end{pmatrix}$ according to the distribution of $A_1+2A_2+2A_3+A_4 \pmod 3$. Other notations and plotting conventions as in Figs. 1 and 2.

Because the two combinations A_1+A_2 (Fig. 2A) and A_1+2A_2 (Fig. 2B) that define the two genera are linearly independent, their probability distributions can be specified independently; we exploit this in Experiment 2.

The same parameterization strategy can be applied in the other grid directions, yielding a pair of vectors $\vec{\sigma}_{\begin{pmatrix}1\\1\end{pmatrix}}$ and $\vec{\sigma}_{\begin{pmatrix}1\\2\end{pmatrix}}$ for the genera within the family of correlations between pairs of checks that are vertically adjacent, the vectors $\vec{\sigma}_{\begin{pmatrix}1&\\&1\end{pmatrix}}$, $\vec{\sigma}_{\begin{pmatrix}1&\\&2\end{pmatrix}}$ for the genera within the family of correlations in the upper-left to lower-right direction, and $\vec{\sigma}_{\begin{pmatrix}&1\\1&\end{pmatrix}}$, and $\vec{\sigma}_{\begin{pmatrix}&1\\2&\end{pmatrix}}$ for the genera within the family of correlations in the upper-right to lower-left direction. We refer to $\vec{\sigma}_{(1\ \ s)}$ and $\vec{\sigma}_{\begin{pmatrix}1\\s\end{pmatrix}}$ as "cardinal second-order correlations," and to $\vec{\sigma}_{\begin{pmatrix}1&\\&s\end{pmatrix}}$ and $\vec{\sigma}_{\begin{pmatrix}&1\\s&\end{pmatrix}}$ as "diagonal second-order correlations." Each of these eight genera have two degrees of freedom ("species"), corresponding to the triangular domain of the distribution of values for its linear combination. Thus, there are a total of 16 free parameters for the second-order correlations: four families of vectors $\vec{\sigma}_{(1\ \ s)}$, $\vec{\sigma}_{\begin{pmatrix}1\\s\end{pmatrix}}$, $\vec{\sigma}_{\begin{pmatrix}1&\\&s\end{pmatrix}}$, $\vec{\sigma}_{\begin{pmatrix}&1\\s&\end{pmatrix}}$, each with two genera ($s=1$ and $s=2$), and these eight vectors each occupy a triangular domain. This is a substantial expansion compared to the black-and-white case, where there were a total of 4 free parameters (β_{-}, $\beta_{|}$, $\beta_{\backslash}$, and $\beta_{/}$; see Table 1).

We also mention the correspondence to the notation of [23] for second-order statistics: the subscripts 1 or 2 of $\vec{\sigma}$, used here, correspond to the subscripts $+$ and $-$ of β in [23]. The numerical notation used here generalizes more readily to multiple gray levels.

Third- and fourth-order statistics

The analogous approach provides a parameterization of third- and fourth-order correlations. For example, there is a family of third-order statistics corresponding to the correlations among the three checks in the $\ulcorner$-shaped region $\begin{pmatrix}A_1&A_2\\A_3&\end{pmatrix}$. This family is subdivided into four genera, corresponding to the distributions of the four sums $A_1+A_2+A_3\ (\mathrm{mod}\,3)$, $A_1+A_2+2A_3\ (\mathrm{mod}\,3)$, $A_1+2A_2+A_3\ (\mathrm{mod}\,3)$, and $A_1+2A_2+2A_3\ (\mathrm{mod}\,3)$, which are linearly independent. As in the second-order case, each of these genera is a triangular domain, whose coordinates indicate the probability that the sum $A_1+s_2A_2+s_3A_3$ is 0, 1, or 2.

Fig. 3A shows the domain parameterized by $\vec{\sigma}_{\begin{pmatrix}1&1\\1&\end{pmatrix}}$, the vector that specifies the distribution of the sum $A_1+A_2+A_3\ (\mathrm{mod}\,3)$. At the bottom vertex of the triangle, $\vec{\sigma}_{\begin{pmatrix}1&1\\1&\end{pmatrix}}=(1,0,0)$, so $A_1+A_2+A_3=0\ (\mathrm{mod}\,3)$. Since this relationship holds whenever the three A_k's are equal, the resulting texture contains $\ulcorner$-shaped regions uniformly black, gray, or white. At the other two vertices, $A_1 + A_2 + A_3 = 1 \ (\bmod 3)$ or $A_1 + A_2 + A_3 = 2 \ (\bmod 3)$. Every Γ-shaped region therefore must contain at least two different luminance levels. Since there are four possible orientations of a Γ-shaped region, there are four such families of third-order statistics (each with the analogous four genera, and two degrees of freedom in each genus), for a total of 32 independent third-order statistics.

At fourth-order, there is a single family, corresponding to the entire 2×2 neighborhood. Fig. 3B shows an example domain, corresponding to the genus $\vec{\sigma}_{\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)}$, which specifies the distribution of $A_1 + 2A_2 + 2A_3 + A_4 \ (\bmod 3)$. At the bottom vertex of this domain, where $\vec{\sigma}_{\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)} = (1,0,0)$, the texture has uniform 2×2 regions of all luminance levels. This is because $A_1 + 2A_2 + 2A_3 + A_4 = 0 \ (\bmod 3)$ is equivalent to $A_1 + A_4 = A_2 + A_3 = 0 \ (\bmod 3)$, which holds for any constant value of the A_k. In total, there are 16 independent fourth-order statistics, corresponding to the eight genera, $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&2\\1&1\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\2&1\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\1&2\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&2\\1&2\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\2&2\end{smallmatrix}\right)}$, and $\vec{\sigma}_{\left(\begin{smallmatrix}1&2\\2&2\end{smallmatrix}\right)}$, each of which is a triangular domain with two free parameters. For further details and correspondences to the black-and-white case, see Table 1 and the Supplemental Document ("Specification and construction of textures").

Experiment 1: Individual texture statistics, three luminance levels

Experiment 1 quantifies sensitivity within each of the triangular domains (genera) described above: one first-order domain (Fig. 1), eight second-order domains (two examples shown in Fig. 2), 16 third-order domains (an example shown in Fig. 3A), and eight fourth-order domains (an example shown in Fig. 3B). The burden of studying these 33 domains, each containing two degrees of freedom, may be reduced by recognizing that many of them are interrelated by spatial symmetries. For example, exchanging horizontal and vertical axes interconverts $\vec{\sigma}_{(1\ \ s)}$ and $\vec{\sigma}_{\left(\begin{smallmatrix}1\\s\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&s_2\\s_3&\end{smallmatrix}\right)}$ and $\vec{\sigma}_{\left(\begin{smallmatrix}1&s_3\\s_2&\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&s_2\\s_3&s_4\end{smallmatrix}\right)}$ and $\vec{\sigma}_{\left(\begin{smallmatrix}1&s_3\\s_2&s_4\end{smallmatrix}\right)}$, etc. Additional symmetries include mirror-flips and 90-deg rotations. Previous work with black-and-white textures showed that statistics related by these symmetries had the same thresholds [13], and, in preliminary experiments, we verified that this equivalence held for the cardinal second-order correlations in the textures with three luminance levels. We therefore limited our analysis to 12 domains, from which all other domains could be obtained via a symmetry operation. These domains were: the first-order domain $\vec{\sigma}_{(1)}$; the second-order domains $\vec{\sigma}_{(1\ \ 1)}$, $\vec{\sigma}_{(1\ \ 2)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&\\&1\end{smallmatrix}\right)}$, and $\vec{\sigma}_{\left(\begin{smallmatrix}1&\\&2\end{smallmatrix}\right)}$; the third-order domains $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\1&\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\2&\end{smallmatrix}\right)}$, and $\vec{\sigma}_{\left(\begin{smallmatrix}1&2\\2&\end{smallmatrix}\right)}$; and the fourth-order domains $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\1&1\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\1&2\end{smallmatrix}\right)}$, $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\2&2\end{smallmatrix}\right)}$, and $\vec{\sigma}_{\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)}$.

To measure sensitivity to these statistics, we determined psychophysical thresholds in a standard texture-segmentation task ([17], described below) via a method of constant stimuli. For each threshold measurement, stimuli were defined by equally-spaced points lying along 12 rays in the triangular domain. Each ray began at the origin of the domain (the random texture) and extended either towards a vertex, or to points that were equally spaced along the edges of the domain. The distances of the endpoints from the origin were chosen based on pilot experiments so that the texture contrasts would capture the transition between subthreshold and suprathreshold performance. For the first-order statistics, an additional set of 12 rays were interleaved to better delineate the threshold behavior. These specifics are detailed in the Supplemental Document, Figure S1.

For construction of psychophysical curves and quantification of thresholds, the texture contrast c is defined as the distance from the origin (i.e., the centroid of the domain), scaled so that the vertices of each domain have a texture contrast of 1.

Experiment 2: Pairs of texture statistics, three luminance levels

Experiment 2 quantifies sensitivity to combinations of image statistics drawn from different triangular domains (genera). We focused on combinations of cardinal second-order statistics, as sensitivity to these statistics was high, and included interactions between image statistics that specify correlations in the same spatial orientation (i.e., between two image statistics from the $\vec{\sigma}_{(1\ \ s)}$-family) as well as interactions between image statistics that specify correlations in different orientations (i.e., between the $\vec{\sigma}_{(1\ \ s)}$-family and the $\vec{\sigma}_{\left(\begin{smallmatrix}1\\ s\end{smallmatrix}\right)}$-family).

Fig. 4A,B shows stimuli that probe interactions between correlations in the $\vec{\sigma}_{(1\ \ s)}$-family, but drawn from different genera: $\vec{\sigma}_{(1\ \ 1)}$ (along the abscissae of the panels) and $\vec{\sigma}_{(1\ \ 2)}$ (along the ordinates of the panels). The panels differ in terms of the species of the $\vec{\sigma}_{(1\ \ 1)}$ genus that lies along the abscissa: in Fig. 4A, it is in the direction of the $(1,0,0)$-vertex of the $\vec{\sigma}_{(1\ \ 1)}$-domain; in Fig. 4B, it is in the direction of the $(0,0,1)$-vertex. In both cases, the ordinate is in the direction of the $(1,0,0)$-vertex of the $\vec{\sigma}_{(1\ \ 2)}$-domain. Not all combinations of coordinates are represented in these panels, because extreme values of one coordinate limit values of the other – but these limits were beyond the range needed to determine thresholds.

Fig. 4C,D shows stimuli that probe interactions in different spatial directions: $\vec{\sigma}_{\left(\begin{smallmatrix}1\\ 2\end{smallmatrix}\right)}$, along the ordinate, and $\vec{\sigma}_{(1\ \ 2)}$, along the abscissa. In both cases, the ordinate is in the direction of the $(1,0,0)$-vertex of the $\vec{\sigma}_{\left(\begin{smallmatrix}1\\ 2\end{smallmatrix}\right)}$-domain. In Fig. 4C, the abscissa is in the direction of the $(1,0,0)$-vertex in the $\vec{\sigma}_{(1\ \ 2)}$-domain; in Fig. 4D, the abscissa is in the direction of the $(0,1,0)$-vertex of that domain.

Experiments were organized into four groups. Group I examined interactions between different statistics with the same family ($\vec{\sigma}_{(1\ \ 1)}$ and $\vec{\sigma}_{(1\ \ 2)}$, as in Fig. 4A,B); the other groups probed interactions between statistics from different families, describing correlations

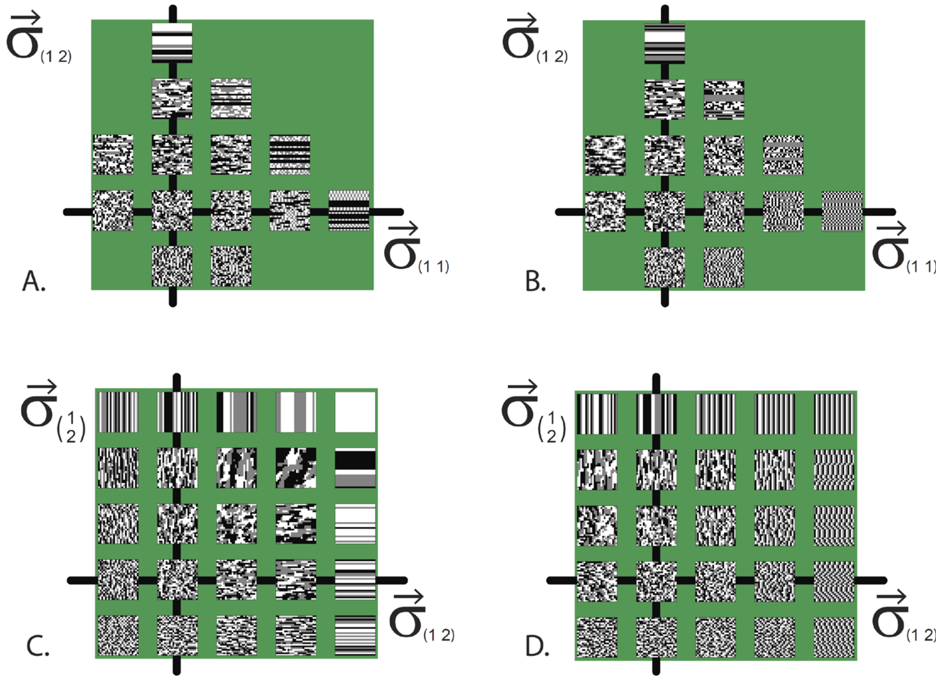

Fig. 4. The domain generated by specifying a pair of cardinal second-order statistics. In each case, the random texture is at the origin, indicated by the intersection of the two solid black lines. Panels A and B: The statistics are $\vec{\sigma}_{(1\ \ 1)}$ and $\vec{\sigma}_{(1\ \ 2)}$, both specifying horizontal correlations. In A, the abscissa indicates values of $\vec{\sigma}_{(1\ \ 1)}$, ranging from $(\frac{1}{9},\frac{4}{9},\frac{4}{9})$ to $(1,0,0)$; this corresponds to the ray pointing towards the lower vertex of Fig. 2A. The ordinate indicates values of $\vec{\sigma}_{(1\ \ 2)}$ over the same range; this corresponds to the ray pointing towards the lower vertex of Fig. 2B. Steps along each axis are equal to $(\frac{2}{9},-\frac{1}{9},-\frac{1}{9})$. In B, the ordinate is the same as in A, but the abscissa now indicates values of $\vec{\sigma}_{(1\ \ 1)}$ ranging from $(\frac{4}{9},\frac{4}{9},\frac{1}{9})$ to $(0,0,1)$, corresponding to the ray pointing towards the upper vertex in Fig. 2A. Here, abscissa steps are equal to $(-\frac{1}{9},-\frac{1}{9},\frac{2}{9})$. Panels C and D: The statistics are $\vec{\sigma}_{(1\ \ 2)}$ and $\vec{\sigma}_{\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)}$, specifying horizontal and vertical correlations, respectively. In C, the abscissa indicates values of $\vec{\sigma}_{(1\ \ 2)}$, ranging from $(\frac{1}{9},\frac{4}{9},\frac{4}{9})$ to $(1,0,0)$; this corresponds to the ray pointing towards the lower vertex of Fig 2B. The ordinate indicates values of $\vec{\sigma}_{\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)}$ over the same range; this corresponds the same kind of correlation, but now in the vertical direction. Steps along each axis are equal to $(\frac{2}{9},-\frac{1}{9},-\frac{1}{9})$. In D, the ordinate is the same as in C but the abscissa now indicates values of $\vec{\sigma}_{(1\ \ 2)}$ ranging from $(\frac{4}{9},\frac{1}{9},\frac{4}{9})$ to $(0,1,0)$, corresponding to the ray pointing towards the right vertex in Fig. 2B. Here, abscissa steps are equal to $(-\frac{1}{9},\frac{2}{9},-\frac{1}{9})$. In all panels, the coordinates at the origin are equal to $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$, corresponding to the random texture.

in orthogonal directions, $\vec{\sigma}_{(1\ \ s)}$ and $\vec{\sigma}_{\left(\begin{smallmatrix}1\\s'\end{smallmatrix}\right)}$: with $(s,s')=(1,1)$ in group II, $(s,s')=(1,2)$

in group III, and $(s,s')=(2,2)$ in group IV (the latter shown in Fig. 4C,D). These domains included 22 pairs of statistics, each sampled along rays in 12 equally-spaced directions. As was the case for Experiment 1, the positions of the three points sampled along each ray were determined by pilot studies to ensure that they would be effective for measuring thresholds; further details are provided in the Supplemental Document, Figure S2 and Table S1. Also as in Experiment 1, texture contrast c is defined as the distance from the origin, scaled so that the vertices of each domain have a texture contrast of $c=1$.

2.2 Stimuli for experiment 3

The last set of experiments makes use of textures with up to 11 luminance levels. We focused on second-order statistics that provided tests of the computational model complementary to the data of Experiments 1 and 2. Specifically, we selected members of the second-order families $\vec{\sigma}_{(1\;\;s)}$ and $\vec{\sigma}_{\binom{1}{s'}}$ that specify progressively smoother gradients as further gray levels were added, and a contrasting set of statistics that does not specify gradients.

The textures that probe these statistics are shown in Fig. 5, for the number of gray levels (3, 4, 5, 7, and 11) used in these experiments. Within each of the "gradient" stimuli (the individual patches in Fig. 5A), luminances tend to increase gradually in one direction (here, left-to-right), and then reset abruptly from white to black. This progression is clearest for the examples with maximum correlation strength ($c=1$). In the "streaks" (Fig. 5B), luminances of adjacent checks (here, horizontal) tend to match, leading to elongated streaks as correlation strength increases. In all cases, a correlation strength of 0 corresponds to the random texture, and the maximum correlation strength is 1.

The statistics underlying these textures are specified by an extension of the formalism used above for three-luminance textures (see Table 1 and Supplemental Document). The G gray levels are designated by 0, 1, …, $G-1$, where, by convention, we take 0 to indicate a gray-level of black, and $G-1$ to indicate a gray-level of white. A horizontal second-order image statistic $\vec{\sigma}_{(1\;\;s)}$ is specified by a vector $(v_0,v_1,\ldots,v_{G-1})$ of length G, each entry v_h is the probability that $A_1+sA_2=h \pmod{G}$. Vertical second-order image statistics $\vec{\sigma}_{\binom{1}{s}}$ are specified analogously. All coordinates are non-negative and must sum to 1, and the random texture corresponds to a vector $(v_0,v_1,\ldots,v_{G-1})$ with all entries equal to $1/G$. Note that G, the number of gray levels, and c, the texture contrast, are independent. The G gray levels always include black and white and $G-2$ equally-spaced intermediate values, and each gray level occurs in $1/G$ of the checks. Independently, c, the texture contrast, indicates the departure of the spatial arrangement from randomness.

A left-to-right gradient texture can be represented in these coordinates as follows. In a left-to-right gradient, the probability that $A_2=A_1+1$ is increased. This is equivalent to an increase in the probability that $A_1-A_2=-1 \pmod{G}$, i.e., that $A_1+(G-1)A_2=G-1 \pmod{G}$. Thus, the relevant image statistic is $\vec{\sigma}_{(1\;\;G-1)}$, and its final (($G-1$)th) entry captures this increase. Since the bias increases linearly with correlation strength, the parameterization of this gamut of textures is given by

$$\vec{\sigma}_{(1\;\;G-1)}=(1-c)\left(\tfrac{1}{G},\tfrac{1}{G},\ldots,\tfrac{1}{G},\tfrac{1}{G}\right)+c\left(0,0,\ldots,0,1\right). \quad (1)$$

At maximum correlation strength ($c = 1$), $G-1$ is the only permissible value of $A_1 + (G-1)A_2 \pmod{G}$, so that $A_2 = A_1 + 1 \pmod{G}$ -- yielding a strict gradient with luminances increasing to the right. At zero correlation strength ($c = 0$), all values are equally likely, yielding a random texture. These conventions are consistent with that of Experiments 1 and 2, in which $c = 0$ corresponds to a random texture and $c = 1$ corresponds to a maximally-structured texture at a vertex of the domain.

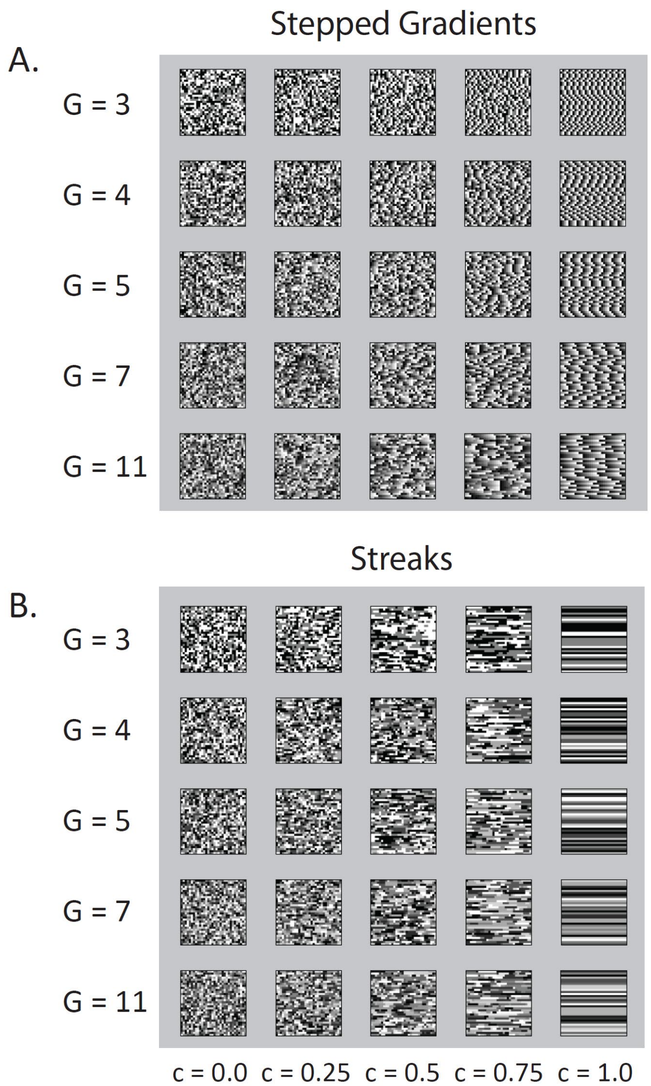

Fig. 5. Examples of textures used in Experiment 3: gradients (A) and streaks (B). Number of gray levels indicated by G. Gradient textures have a directionality – in the examples shown here, from left to right. In the direction of the gradient, the choice of luminance in each check is biased towards a stepwise increase from black to white, followed by an abrupt decrease to black. This is most evident in the examples with maximal correlation strength (right end of each row, $c = 1$): here, luminances in each row of checks progressively increase from black to white and then reset to black; the phase of each row is random. Streak textures (B) have an orientation – in the examples shown here, horizontal -- but not a directionality. Along the specified orientation, the choice of luminance in each check is biased to match its neighbor. At maximal correlation strength (right end of each row, $c = 1$), this results in rows of checks whose luminance is constant. In all cases, a correlation strength of 0 corresponds to the random texture.

Similarly, a leftward gradient texture is specified by increasing the probability that $A_2 = A_1 - 1$, i.e., that $A_1 - A_2 = 1 \pmod G$. So these textures are parameterized by

$$\vec{\sigma}_{(1\quad G-1)} = (1-c)\left(\tfrac{1}{G},\tfrac{1}{G},\ldots,\tfrac{1}{G},\tfrac{1}{G}\right) + c\left(0,1,\ldots,0,0\right). \tag{2}$$

Streaks are created by increasing the probability that $A_2 = A_1$, i.e., that $A_1 - A_2 = 0 \pmod G$. Thus, streaks are parameterized by

$$\vec{\sigma}_{(1\quad G-1)} = (1-c)\left(\tfrac{1}{G},\tfrac{1}{G},\ldots,\tfrac{1}{G},\tfrac{1}{G}\right) + c\left(1,0,\ldots,0,0\right). \tag{3}$$

Downward and upward gradients and vertical streaks are parameterized in a similar fashion, with $\vec{\sigma}_{\left(\begin{smallmatrix}1\\ G-1\end{smallmatrix}\right)}$ replacing $\vec{\sigma}_{(1\quad G-1)}$.

As in Experiments 1 and 2, $c = 0$ corresponds to the random texture and $c = 1$ corresponds to maximal correlation strength – periodic ramps for the gradient texture (eqs. (1) and (2)), and unbroken lines of constant luminance for the streak texture (eq. (3)).

2.3 Subjects

Studies were conducted in 10 normal subjects (3 male, 7 female), ages 21 to 55. Two of the subjects (MC and SR) were experienced psychophysical observers. MC, SR, and JB are authors; LE assisted with the studies; the other observers were naïve to the purposes of the experiment. All subjects had visual acuities (corrected if necessary) of 20/20 or better.

Experiment 1 was conducted in five subjects (MC, SR, NM, WC, ZA) for all first- and second-order statistics, and for all third- and fourth-order statistics for which thresholds could be obtained. For Experiment 2, group I was conducted in MC and WC, group II in MC and ZA, group III in MC and JB, and group IV in MC, WC, ZA, and JB. Experiment 3 was conducted in six subjects (MC, IL, LE, YCL, EFV, PJ).

This work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), following approval of the Institutional Review Board of Weill Cornell, and consent of the individual subjects.

2.4 Segmentation task

For the three experiments, segmentation thresholds were measured in a four-alternative task adapted from the one developed by Chubb et al., [17] and identical to what was used in related previous studies[13, 27, 28]. We describe it below for the reader's convenience, along with the initial analysis steps.

All stimuli consisted of 64×64 arrays of checks; each contained an embedded 16×64 rectangular target whose outer edge was 8 checks from one of the four sides of the array. Target and background regions were filled either with a structured texture drawn from one of the domains described above, or a contrasting texture. The contrasting texture was fully random, i.e., each check was independently colored with one of G equally-spaced luminance values from black to white, each with probability $1/G$ ($G = 3$ in Experiments 1 and 2; $G \in \{3,4,5,7,11\}$ in Experiment 3). To ensure that the subject performed the task by identifying a texture boundary, rather than a texture gradient [31], half of the trials had a structured target on a random background and half had a random target on a structured background. Examples of both kinds of stimuli (structured target on random background, random target on structured background) are shown in Fig. 6A, and the four alternative target positions are shown in Fig. 6B.

As previous work showed no consistent threshold difference between these conditions, we pooled data across this randomization. We also found no consistent difference between

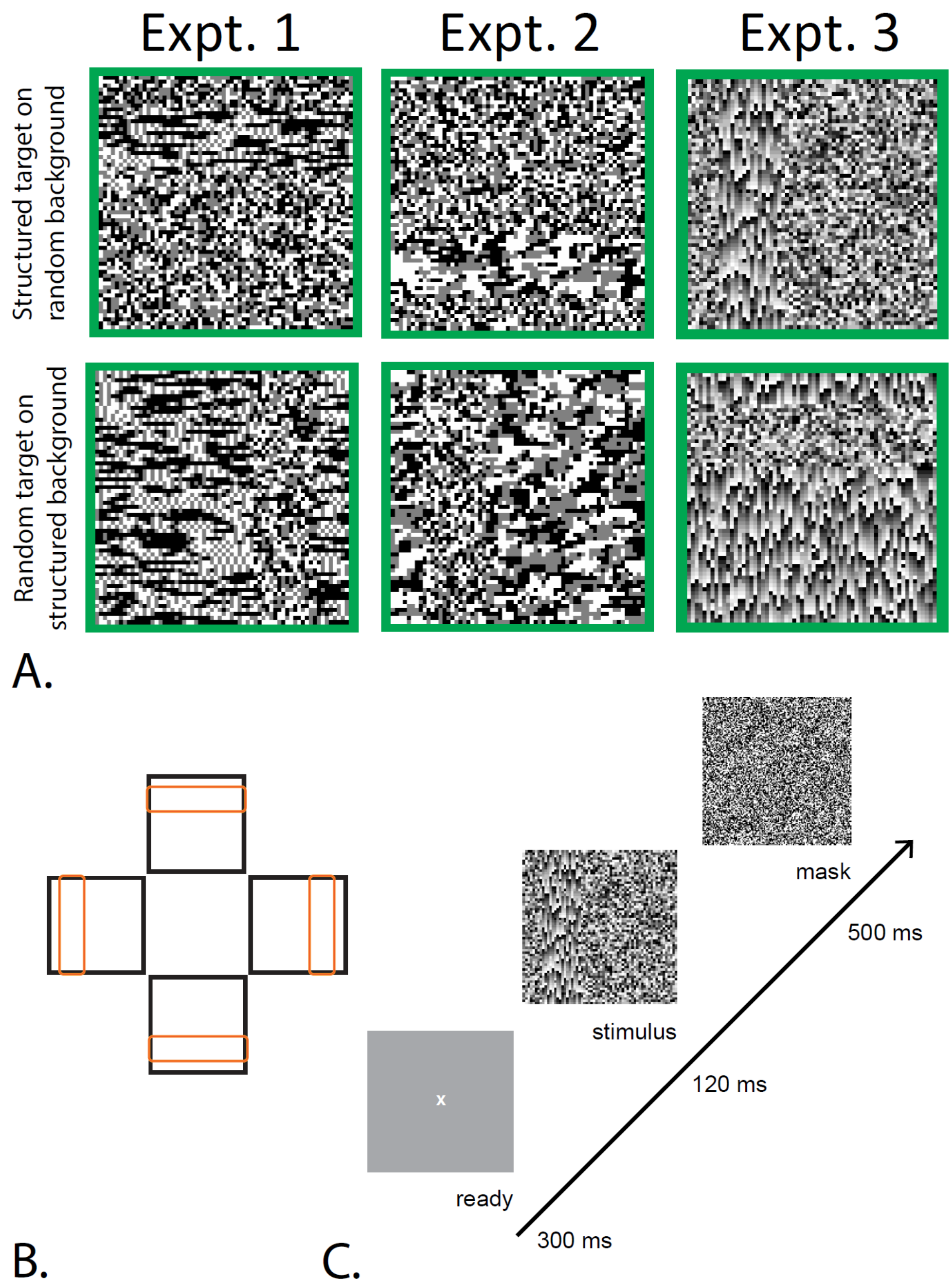

Fig. 6. A: Stimulus examples for the three experiments. Stimulus parameters: Experiment 1, from domain of Fig. 2A, with $\vec{\sigma}_{(1\ \ 1)} = [0.75, 0.25, 0]$ (texture contrast $c = 0.66$); Experiment 2, from domain of Fig. 4C, $\vec{\sigma}_{(1\ \ 2)} = [0.6218\ 0.1891\ 0.1891]$ and $\vec{\sigma}_{\binom{1}{2}} = [0.5\ 0.25\ 0.25]$ (texture contrast $c = 0.5$); Experiment 3, from Fig. 5A, with $G = 11$ gray levels and texture contrast $c = 0.8$. All of these texture contrasts are suprathreshold. B: The four alternative target positions. C: Trial timeline.

thresholds for horizontal vs. vertical stimuli in Experiment 3, and therefore pooled across these conditions. Note also that, while this task has a "global" component in the sense that evidence can be pooled across the entire stimulus, this global aspect is constant across all stimuli; the limiting factor is the information contained in local correlations, which is stimulus-dependent.

In Experiment 1, each test session explored a triangular domain specified by a texture-statistic genus; examples of these domains are shown in Fig. 1 (first-order), Fig. 2 (second-order), and Fig. 3 (third- and fourth-order). Second-, third-, and fourth-order domains were sampled along 12 rays (Fig. S1A,B); the first-order domain was sampled along 24 rays (Fig. S1C) in separate sessions of 12 rays each. Three texture contrasts were chosen along each ray to span the range from near-chance performance to near-perfect performance in pilot experiments, or, if performance did not achieve near-perfect performance, at texture contrasts (c) of $1/3$. $2/3$, and 1. A single test session contained 8 examples of stimuli specified by all three texture contrasts on the 12 rays; these 8 examples included each of the four target positions, and in both target-structured and background-structured conditions, yielding $3\times12\times4\times2=288$ unique trials, presented in random order. We collected responses to 15 such 288-trial blocks from each subject, yielding 120 judgments for each of the three contrast levels on each ray. For third- and fourth-order statistics, results from two subjects (MC, SR) showed that sensitivity was largely restricted to a subset of three rays; in these domains, the other subjects (NM, WC, ZA) were tested with only these three rays. In these cases, blocks contained 32 examples of each contrast level on each ray and 4 such blocks were obtained, yielding 128 judgments for each contrast level on each ray.

Experiment 2 was organized similarly, with each test session devoted to a domain specified by a pair of second-order texture statistics; examples are shown in Fig. 4 and the sampling strategy is given in Table S1 and Fig. S2.

In Experiment 3, each test session consisted of stimuli with a fixed number of gray levels (3, 4, 5, 7, or 11), and included both gradient stimuli (eqs. (1) and (2)) and streak stimuli (eq (3)). To cover the range of performance, five texture contrasts were used: $c\in\{0.2,0.3,0.45,0.6,0.8\}$ for the gradient stimuli, and 2/3 of these values for the streak stimuli. There were 6 kinds of stimuli: gradients in each of the four cardinal directions contrasted with the random texture, and streaks in horizontal and vertical orientations contrasted with the random texture. As in Experiments 1 and 2, targets appeared in each of four possible positions, and the textures used to render the target and background were swapped in half of the trials. Thus, there were $5\times6\times4\times2=240$ unique trials, presented in random order. We collected responses to 12 such blocks from each subject, yielding 96 judgments for each of the five contrast levels and the six kinds of stimuli.

We collected data from six subjects (MC, IL, LE, YCL, EFV, PJ) for 3, 5, and 11 gray levels and in four of these (MC, YCL, EFV, and PJ) for 4 and 7 gray levels.

2.5 Procedure

The procedure for the three experiments was similar to that of previous studies [13, 27, 28] and is summarized here. A Cambridge Research ViSaGe system, running custom Delphi software produced the stimuli and collected responses. Stimuli were displayed on an LCD monitor (mean luminance of 23 cd/m^2, refresh rate 60 Hz), beginning 300 ms after the subject pressed a "ready" button. Stimuli had a duration of 120 ms, and were followed by a 500-ms mask consisting of checks that were half the size of the stimulus checks, randomly filled with the luminance levels used in the experiment (see Fig. 6C for the timeline). The display size was 15×15 deg (64×64 checks, 14.8 min each, each check consisting of 10x10 monitor pixels); viewing was binocular at 100 cm, and contrast was 1. Note that the checks were sufficiently large so that, even at the edges of the display, they were plainly visible[32], and previous work with black-and-white textures showed that thresholds are approximately scale-invariant at and below this check size [13]. ViSaGe software and its photometer was used to linearize the monitor's output via a look-up table, which was recalibrated prior to each experimental session. Thus, the luminance levels used ranged from 0 cd/m^2 (black checks) to 46 cd/m^2 (white checks). The gray checks in Experiments 1 and 2 were 23 cd/m^2; the gray checks in Experiment 3 had luminance levels equally spaced between 0 and 46 cd/m^2 – for example, for $G = 5$, the luminance levels were 0, 11.5, 23, 34.5, and 46 cd/m^2.

Subjects were informed that on every trial, a target would be present, and was equally likely to be in any of four positions (top, right, bottom, left), which they were to indicate by pressing the corresponding button on a four-button response box. They were asked to fixate centrally and not attempt to scan the stimulus. Trials were self-paced, triggered by a separate button-press. Inexperienced subjects received practice of approximately two hours to become accustomed to the brief stimulus presentation time and to practice maintaining central fixation without scanning. During practice, but not during data collection, subjects received auditory feedback for incorrect responses.

2.6 Analysis

For each stimulus type (i.e., for each ray in the texture domains of Experiments 1 and 2, and for each kind of gradient or streak in Experiment 3), we determined the texture contrast threshold for segregation, via a procedure similar to that used in previous studies [13, 26, 27], as summarized here. First, for each set of responses to a given stimulus type, we found the maximum-likelihood fit of a Weibull function to the observed fraction correct (FC):

$$FC(c) = \frac{1}{4} + \frac{3}{4}\left(1 - 2^{-(c/a_r)^{b_r}}\right). \tag{4}$$

As above, c is the texture contrast, defined as the distance to the fully random texture (the centroid), normalized by the distance from the vertex to the centroid. a_r is the fitted threshold (i.e., the value of c at which FC=0.625, halfway between chance (0.25), and perfect (1.0)), and b_r is the Weibull shape parameter. As previously reported [13, 27], the shape parameter b_r typically had similar values across rays, with overlapping confidence limits that usually included the range 2.2 to 2.7. Since our focus is on determining the thresholds, we then refit the data from each experiment by a set of Weibull functions that shared a common shape parameter b, while allowing the threshold parameter a_r to vary freely across rays. This procedure reduced the number of free parameters without altering the quality of the fit to Weibull functions. 95% confidence intervals were determined via 1000-sample bootstraps. Note that this procedure could yield an estimated threshold $a_r > 1$, i.e., beyond the boundary of the texture domain, if performance was above chance but never reached a FC of 0.625.

Sensitivity was defined as 1/threshold, with corresponding confidence intervals. Across-subject averages of sensitivities or thresholds are computed as the geometric means, and statistics are computed on the logarithms of the raw values. All calculations were carried out with in-house MATLAB (MathWorks, Natick, MA) software, which was also used to synthesize the stimuli described above.

3. Model

Here we describe a computational model for discrimination thresholds for textures that contain multiple gray levels and spatial correlations (Fig. 7). As a starting point, we used two complementary sets of psychophysical studies: studies of textures with multiple gray levels but without spatial correlation ("IID textures"), and studies with spatial correlation but only black and white checks. These studies were carried out with different paradigms, in separate labs, and with separate subjects. The model described here is fully constrained by these studies and makes explicit predictions for discrimination thresholds for textures that include multiple gray levels and spatial correlations.

In overview, the model (Fig. 7) is as follows. The first stage of the model accounts for sensitivity to IID textures by recasting the mechanisms proposed by the studies of Chubb and colleagues [14, 16, 17, 33] as stochastic thresholds, rather than gray-level sensitivities. This stage yields a set of internal representations, one for each of the original Chubb mechanisms. The second stage of the model then processes the local correlations within these internal representations. The computations used to do this are the same as those deduced in our previous studies [13, 18, 27, 28] that focused on black-and-white textures.

We note that, while we describe the model's computations in terms of the texture coordinates introduced above, the model operates directly on the visual input. Thus, it makes predictions that are independent of the coordinates used to parameterize the textures, it treats all orders of correlation together, and it is not restricted to the textures that lie within the space we consider.

3.1 First stage: sensitivity to gray-level distribution

Chubb and colleagues [14, 17, 33] showed that discrimination of IID textures could be accounted for by three "dimensions:" one dimension approximating the mean luminance, a second dimension approximating variance, and a third dimension signaling the fraction of very dark checks ("blackshot"). Coordinates along dimension m were linear functions of the histogram distribution:

$$c_m = \sum_i D_m(x_i) g(x_i), \tag{5}$$

where the sum ranges over the gray levels in the texture, $g(x)$ is the frequency with which gray level x occurs, and $D_m(x)$ is the extent to which a gray level x contributes to mechanism m. IID textures that shared the same coordinates (c_1, c_2, c_3) were indistinguishable, even if their gray-level distributions were disparate. Using an asymmetric search task, they later [16] showed that these three dimensions derived from the activations of four underlying mechanisms, which were also linear functions of the histogram distribution:

$$a_m = \sum_i F_m(x_i) g(x_i), \tag{6}$$

These four mechanisms are necessarily linearly dependent, since they are constrained to yield the three dimensions of eq. (5) above. For textures with nine equally-spaced gray values $\{0, 1/8, ..., 7/8, 1\}$, [16] determined consensus values of the linear functions of $F_m(x_i)$ across three subjects, along with the relative weightings with which each subject used these mechanisms. These data were kindly provided by C. Chubb and are given in Table S2. The correspondence to the nomenclature of [16] is as follows: F_1 and F_2 correspond to the two complementary quasilinear mechanisms (their $F_{*,3}$ and $F_{*,4}$); F_3 corresponds to the blackshot-like mechanism (their $F_{*,1}$), and F_4 corresponds to the mechanism sensitive to midrange grays (their $F_{*,2}$).

To apply these data to general gray-level distributions, we interpolated these values via a cubic spline. Thus, for a texture in which $g(x)\Delta x$ is the fraction of checks with gray levels between x and $x+\Delta x$, the "activation" produced in mechanism m ($m \in \{1,2,3,4\}$) is given by

$$a_m = \int_0^1 F_m(x) g(x) dx . \tag{7}$$

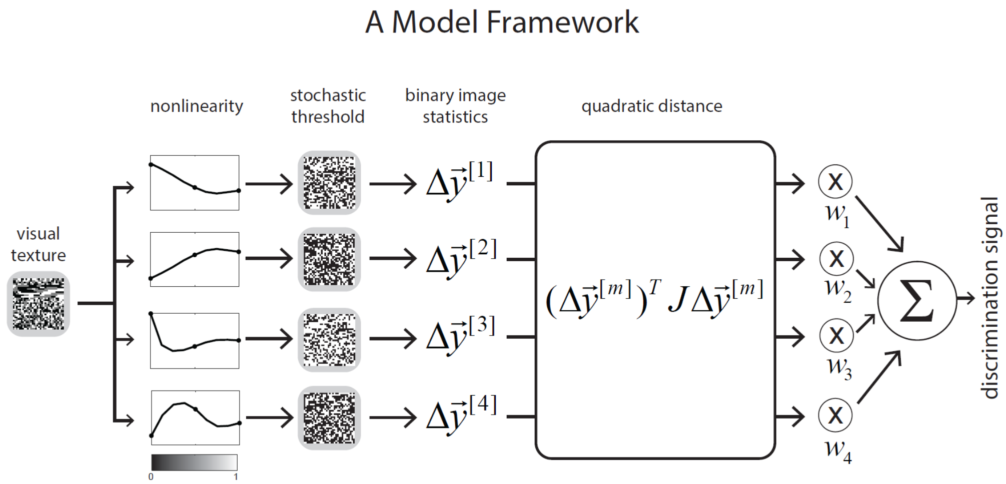

Fig. 7. A model for discrimination of textures with multiple gray levels and spatial correlations, illustrating how it acts on the visual stimuli used here. The four curves labeled "nonlinearity" show the mechanisms F_m^{prob} (eq. (8)). For further details, see text.

In our model, we recast each Silva and Chubb mechanism m as a probabilistic conversion to an internal representation I_m of the original texture. Specifically, we interpret $F_m(x)$ as a nonlinear function of the gray level, whose value at each location in the texture determines the probability that the original check is internally represented in the "high" state (designated 1), vs. the "low" state (designated 0). The probability that a check of gray-level x is converted to 1 by mechanism m is given by

$$F_m^{prob}(x) = \frac{1}{2}\left(1 + \frac{F_m(x)}{\max_{m,x} |F|}\right). \tag{8}$$

This remaps the zero-centered $F_m(x)$'s to quantities $F_m^{prob}(x)$ that range from 0 to 1, as shown by the nonlinearities in Fig. 7. We postulate that this stochastic conversion from gray-level to a binary representation is independent at each check and across the mechanisms.

In this re-interpretation, the spatial average $\langle I_m \rangle$ of the internal representation of a texture with luminance distribution $g(x)$ corresponds to the activation produced by the mechanism in the original formulation, other than a fixed offset and proportionality constant:

$$\begin{aligned}\langle I_m \rangle = \langle F_m^{prob}(x) \rangle = \int_0^1 F_m^{prob}(x) g(x) dx = \frac{1}{2}\int_0^1 \left(1 + \frac{F_m(x)}{\max_{m,x} |F|} \right) g(x) dx \\ = \frac{1}{2} + \frac{1}{2\max_{m,x} |F|}\int_0^1 F_m(x) g(x) dx = \frac{1}{2}\left(1 + \frac{a_m}{\max_{m,x} |F|} \right)\end{aligned}, \quad (9)$$

where we have used eq. (7) and $\int_0^1 g(x)dx = 1$, since it is a probability distribution. This means that two textures are indistinguishable in the original Chubb model if, for each mechanism, their internal representations in the present model have identical average values.

3.2 Second stage: sensitivity to spatial structure

The influence of the spatial organization of these internal representations is addressed by the second stage of the model. Specifically, we posit that texture discrimination is based on comparing the local statistics of these internal representations, and that the local statistics are compared according to the model [13] for black-and-white textures. That model posited that discrimination of a locally-correlated black-and-white texture from a random texture could be accounted for by 10 local image statistics. These quantities, which correspond to the local image statistics introduced above for $G = 2$ (γ, β_{-}, $\beta_{|}$, $\beta_{\backslash}$, $\beta_{/}$, $\theta_{\lrcorner}$, $\theta_{\llcorner}$, $\theta_{\ulcorner}$, $\theta_{\urcorner}$, α) are here collectively denoted by the column vector $\vec{y} = (y_1, y_2, ..., y_{10})$ to facilitate a compact notation. Sensitivity to these image statistics and their combinations was specified by a 10×10 symmetric matrix Q, with the threshold for discrimination from a random texture given by an ellipsoid,

$$\vec{y}^T Q \vec{y} = \sum_{i,j=1} Q_{ij} y_i y_j = S \quad . \quad (10)$$

Q is constrained not only by cross-diagonal symmetry ($Q_{ij} = Q_{ji}$), but also by the empirical finding that thresholds are unchanged after 90-deg rotations of a texture, and after mirroring a texture in the cardinal axes. This leaves a total of 20 free parameters for Q. [13] determined these parameters (for $S = 1$) in 4 subjects (one of whom, MC, was a subject in the present studies) and validated them with out-of-sample predictions for black-and-white textures. Here, we use the average (arithmetic mean) across subjects (Table S2).

To incorporate this process into a model for discrimination of gray-level textures in a way that ensures consistency with findings for black-and-white textures, we need to consider how the characteristics of the mechanisms in the first stage influence the local image statistics $\vec{y}^{[m]}$ of its internal representation. We first consider how a mechanism transforms the probabilities of gray-level configurations into probabilities of binary configurations, and then the transformation from binary configurations into local image statistics.

The key observation is that, although each of the Silva-Chubb mechanisms depends nonlinearly on gray level, they act linearly on the probabilities of local configurations. That is, a 2×2 region of the stimulus texture with gray-level values $\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$ (with x_i in the range $[0,1]$) will be converted by mechanism m to one of the 16 possible binary representations $\begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$ ($b_i = 0$ or 1). The probability that this block will be converted to $\begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$ via mechanism m is equal to the joint probability that each of the x_i is converted to the corresponding internal representation b_i. Since we posit that these conversions are independent,

$$p^{[m]}\begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = F_m(x_1,b_1)F_m(x_2,b_2)F_m(x_3,b_3)F_m(x_4,b_4),\ where \\ F_m(x,b) = \begin{cases} 1 - F_m^{prob}(x),\ b = 0 \\ F_m^{prob}(x),\ b = 1 \end{cases} . \tag{11}$$

The probability of each binary block type in the internal representation is the sum of contributions from each of gray-level configurations in the original texture:

$$p^{[m]}\begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = \sum_{\vec{x}} F_m(x_1,b_1)F_m(x_2,b_2)F_m(x_3,b_3)F_m(x_4,b_4)p\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}, \tag{12}$$

where the sum is over all G^4 gray-level configurations $\vec{x} = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$. This can be written more compactly as

$$\vec{p}^{[m]} = L_G^{[m]}\vec{p}\,, \tag{13}$$

where $\vec{p}$ is a column vector of the probabilities of the G^4 gray-level configurations indexed by $\vec{x}$, $\vec{p}^{[m]}$ is a column vector of the $2^4 = 16$ binary block probabilities in the internal representation m, and $L_G^{[m]}$ is a $G^4 \times 16$ matrix specified by the multiplier in (12). Note that although $L_G^{[m]}$ is a large matrix, it is entirely specified by the Silva-Chubb model and the set of gray levels, via our probabilistic interpretation.

We next consider how $\vec{p}^{[m]}$, the block probabilities of the internal representation, are captured by the binary image statistics $\vec{y}^{[m]}$. Each image statistic is a linear combination of block probabilities -- for example, $\beta_|$, is the difference between the fraction of 2×1 blocks in which the checks match, and the fraction in which they mismatch. Thus, the transformation from block probabilities to image statistics is linear:

$$\vec{y}^{[m]} = Y\vec{p}^{[m]}, \tag{14}$$

where Y is a 10×16 matrix determined solely by the definition of the image statistics (for details, see [18]), and is given in Table S4. Combining eqs. (13) and (14) yields the transformation from $\vec{p}$, the block probabilities in the original texture, to $\vec{y}^{[m]}$, the image statistics of the binary representation produced by mechanism m:

$$\vec{y}^{[m]} = YL_G^{[m]}\vec{p}\,. \tag{15}$$

According to our model, the threshold to distinguish two textures, characterized by $\vec{p}$ and $\vec{p}'$ respectively, is based on a comparison of the image statistics of their binary representations, $\vec{y}^{[m]}$ and $\vec{y}'^{[m]}$. For distinguishing between a structured black-and-white texture with statistics $\vec{y}$ and a random one ($\vec{y}'=0$), we previously found [13]that thresholds were accounted for by a quadratic function of $\vec{y}$ (eq. (10)). For comparison of two structured textures, we previously found [28] that, in the three coordinate planes tested, thresholds depended primarily on the difference $\vec{y}-\vec{y}'$, and not on the reference texture $\vec{y}'$. That is, the texture discrimination signal for two black-and-white textures given is given by:

$$S=(\vec{y}-\vec{y}')^T Q(\vec{y}-\vec{y}') . \tag{16}$$

Here, we postulate that these findings also apply at the level of the internal binary representations, i.e., that each internal binary representation generates a signal based on a quadratic function of the difference in image statistics $\Delta\vec{y}^{[m]}=\vec{y}^{[m]}-\vec{y}'^{[m]}$. The overall texture discrimination signal S is then a sum of contributions from each of the mechanisms:

$$S=\sum_m w_m(\Delta\vec{y}^{[m]})^T J\Delta\vec{y}^{[m]} , \tag{17}$$

where the weights w_m are taken to be the weights from the Silva-Chubb model (Table S2). The matrix J describes how the image statistics are used within each mechanism. We determine it by the requirement that eq. (17) is consistent with previous studies of black-and-white textures, i.e., eq. (16). Note that this requirement means that J will not be the same as the matrix Q, since Q acted on the statistics of a black-and-white texture, while J acts on the statistics of the internal representation of a texture, after it has been transformed by each mechanism m. The calculation of J is detailed in the Supplemental Document ("Specification of the model's quadratic form"), and the resulting matrix is given in Table S3.

3.3 Model summary

In brief, the proposed model (Fig. 7) specifies a difference signal S that governs the discrimination of two gray-level textures. The model has two stages. In the first stage, several independent mechanisms generate distinct internal representations of each texture, by applying a stochastic threshold that depends nonlinearly on the gray level (eqs. (7) to (9)). This stage of the model ensures that for textures without spatial correlation ("IID textures"), the model reproduces the three-dimensional domain (luminance, contrast, blackshot) of discriminable IID textures identified by Chubb and coworkers [14, 16, 17, 33]. Consistency is guaranteed because the first stage uses the same mechanisms as the Chubb et al. model, so IID textures that are indistinguishable according to the Chubb et al. model produce indistinguishable internal representations in the present model.

The second stage of the model confers sensitivity to spatial structure by comparing the local statistics of these binary representations. The specifics of that comparison (eq. (17)) are determined by the requirement that for black-and-white textures, the findings of [13] are recovered.

Other than the arbitrary value of S (eq. (17)) at which discrimination occurs, the model's parameters are determined by complementary previous studies: discrimination of textures with multiple gray levels but no spatial correlation, and textures with only black and white checks, with local spatial correlations in two dimensions. Note that for all textures, the dependence of the discrimination signal on texture contrast is quadratic, but the proportionality contrast depends on the kinds of correlations that are present in the texture, via the model specification. These texture-dependent proportionality constants determine the predicted relative sensitivities.

3.4 Making model predictions

To determine model predictions for the current experiments, we simulate the images generated by the stimulation generation procedure and determine the texture contrast for which the discrimination signal S reaches a threshold value. Since the specific experimental paradigm (check size, stimulus size, target size, viewing time, etc.) used here is the same as that of [13, 18, 27, 28] the model predicts the experimentally-measured discrimination threshold to be the value of the texture contrast for which $S=1$. This computational procedure was modified for rays in which the predicted threshold was high, since the stimulus generation procedure is limited in the range of texture contrasts that can be attained. In those directions, we determine the texture contrast at which $S=1/16$ rather than $S=1$. Then, recognizing the quadratic dependence of discrimination signal on texture contrast, we convert this texture contrast into a predicted threshold by multiplying it by $1/\sqrt{S}=\sqrt{16}=4$.

We also made predictions from alternate models that had the same structure as Fig. 7, but posited a different set of first-stage mechanisms. One such model had just one first-stage mechanism, with a threshold at mid-gray: it mapped all darker-than-mean checks to 0, all lighter-than-mean checks to 1, and randomly assigned mid-gray checks to 0 or 1. Other models were reduced from the model of Fig. 7 by omitting one or more of the Silva-Chubb mechanisms from the first stage. Since mechanisms F_1 and F_2 are equal and opposite – and this complementarity was an essential feature of the findings of [16], these reduced models always included either both of these mechanisms, or neither. In all cases, these alternate models were implemented by repeating the above calculations with the modified set of mechanisms, including a re-calculation of the matrix J in eq. (17) so that the resulting model's predictions remain consistent with our findings for black-and-white textures [13].

To provide an omnibus measure of model predictions, we computed the fraction of the variance of the psychophysical thresholds that was unexplained by the model predictions, calculated by comparing the sum of the squares of the difference in measured and predicted thresholds, to the sum of the squares of the predicted thresholds, without scaling. For this purpose, psychophysical thresholds were averaged across individuals via the geometric mean, as in previous studies[13]. For some conditions, the model predicted an infinite threshold (i.e., the criterion of $S=1/16$ in eq. (17) was never reached at any texture contrast). For those conditions, we used the largest finite threshold that the model predicted in any other condition. To obtain a comparable measure of intersubject reliability, we computed the fraction of the variance of each subject's thresholds that was not explained by each other subject, across conditions in common, and report the median value of these variance fractions across all subject pairs. These computations were carried out separately for Experiments 1 (first- and second-order statistics only), 2, and 3.

4. Results

4.1 Experiment 1

We used a four-alternative segmentation task (Fig. 6) to determine sensitivity to image statistics in textures that contained three gray levels and spatial correlations. Each set of measurements focused on the correlations within a particular spatial template– e.g., a pair of horizontally-adjacent checks – and within this family of correlations, on a specific type ("genus") of correlations. As detailed in Methods, the genus is defined by constraining the distribution of a specific linear combination of luminance values of the checks in the template, where luminance is denoted by 0 for black, 1 for gray, 2 for white, and the linear combination is computed mod 3. So, for example, for the family of correlations between a pair of horizontally-adjacent checks, the genus specified by $\vec{\sigma}_{(1\ \ 1)}$ constrains the sum $A+B$ of adjacent luminance values, while the genus specified by $\vec{\sigma}_{(1\ \ 2)}$ constrains the sum $A+2B$. Since these sums are computed mod 3, they can have the values 0, 1, or 2, and the

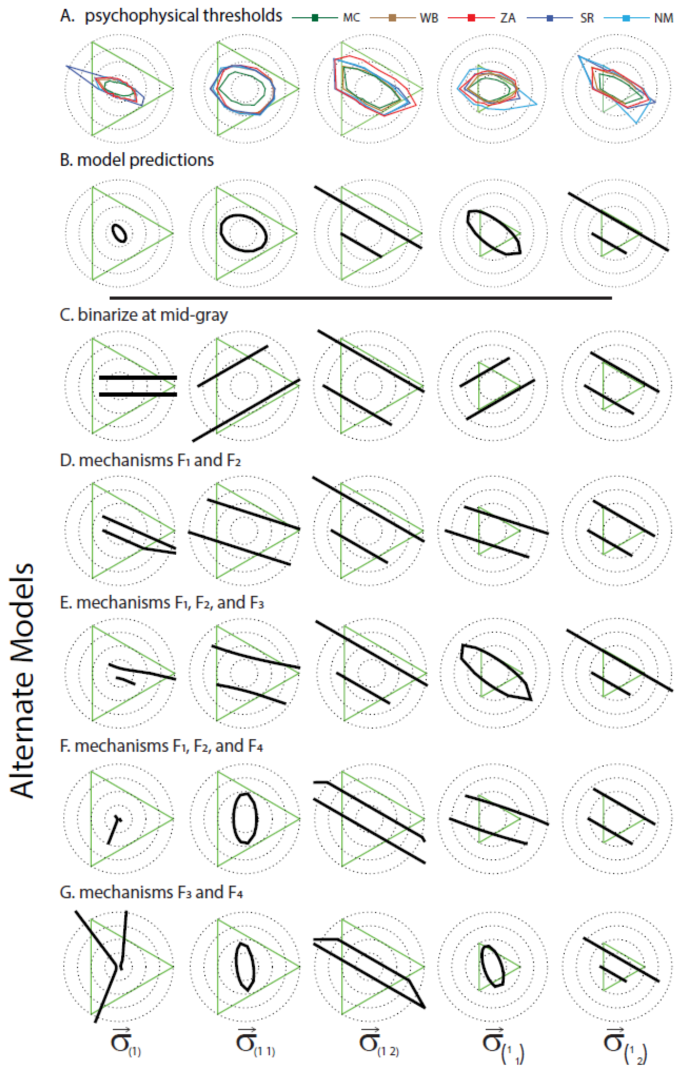

Fig. 8. Psychophysical thresholds (A) and model predictions (B-G) for Experiment 1, first- and second-order statistics. Each triangular domain corresponds to a first-order statistic $\vec{\sigma}_{(1)}$ (Fig.1), a second-order statistic $\vec{\sigma}_{(1\ \ s)}$ involving horizontally-adjacent checks (Fig. 2), or a second-order statistic $\vec{\sigma}_{\begin{pmatrix}1 & \\ & s\end{pmatrix}}$ involving checks that share a corner. Upper row: individual subjects' data. The origin corresponds to a random texture; green triangle corresponds to the boundary of the domain, whose vertices are at $(1,0,0)$. $(0,1,0)$, and $(0,0,1)$. Rings indicate textures of equal correlation strength, with a correlation strength of 1 at the vertices of each domain. Thresholds outside the domain correspond to conditions in which performance was above chance, but did not reach the criterion fraction correct within the stimulus domain. Note that the domains for $\vec{\sigma}_{\begin{pmatrix}1 & \\ & s\end{pmatrix}}$ are plotted on a different scale. For thresholds <1, uncertainties (2SEM) for individual subject thresholds are typically < 10% of the measured thresholds, and are not shown. Model predictions are shown for the full model (B) and for alternate models consisting of a single channel that binarizes at mid-gray(C) or subsets of the Silva-Chubb mechanisms (D-G). For model predictions, isodiscrimination contours are disconnected if predicted thresholds at intermediate directions are >4.

probabilities of these three values describes a triangular domain (see Fig. 2). Similarly, the first-order domain $\vec{\sigma}_{(1)}$ (Fig. 1) is parameterized by the distribution of single-check gray levels, and the third- and fourth-order domains (for example, $\vec{\sigma}_{\begin{pmatrix}1 & 1\\ 1 & \end{pmatrix}}$ and $\vec{\sigma}_{\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}}$, Fig. 3) are parameterized by the distribution of linear combinations of gray levels in 3 and 4 neighboring checks, respectively. In each domain, the centroid is the random texture; the measured thresholds indicate, in multiple directions within each domain, the distance from randomness that is necessary for the statistical structure of the texture to be visually apparent.

Fig. 8A shows the measured thresholds for the first- and second-order domains in five subjects. Thresholds are lowest for the first-order statistics, and, within the second-order statistics, lower for the statistics that describe correlations among horizontally-adjacent checks ($\vec{\sigma}_{(1\ \ s)}$) than for the statistics that describe correlations among checks that share a corner, $\vec{\sigma}_{\begin{pmatrix}1 & \\ & s\end{pmatrix}}$. This pattern holds in all subjects. Additionally, the isodiscrimination contours are approximately elliptical and symmetric about the origin, other than for directions in which thresholds are high and therefore not precisely measurable. Thresholds for correlations among vertically-adjacent checks $\vec{\sigma}_{\begin{pmatrix}1\\ s\end{pmatrix}}$ were not systematically determined, as pilot experiments (as well as our previous studies with black-and-white textures [13]) showed that they were very similar to the corresponding thresholds for horizontally-adjacent checks.

As detailed in Methods, we constructed a model (Fig. 7) for discrimination of spatially-correlated gray-level textures, based on two previously-obtained complementary datasets: (i) studies of discrimination of gray-level textures with no spatial correlation [14, 16, 17, 33], and (ii) studies of discrimination of spatially-correlated black-and-white textures [13, 18, 27, 28]. In brief, the model had two stages: a first stage that analyzed luminance distributions via multiple parallel mechanisms, and produced an internal binary representation along each channel, and a second stage that was sensitive to spatial correlations present within each of these internal representations. The model had no free parameters, as it was fully constrained by the requirement that it accounted for these two previous complementary datasets.

Fig. 8B shows the thresholds predicted by this model. The model approximates the absolute thresholds found experimentally, and fully accounts for the ordering of thresholds among the correlation types. It also accounts for the elliptical shapes of the isodiscrimination contours where thresholds can be reliably determined. However, the model is clearly imperfect. It predicts a greater sensitivity for first-order statistics than we observed, and the axes of the ellipses are inaccurately predicted for $\vec{\sigma}_{(1)}$, $\vec{\sigma}_{(1\ \ 1)}$, and $\vec{\sigma}_{\begin{pmatrix}1 & \\ & 1\end{pmatrix}}$. Note, however, that all model parameters were determined from independent experiments involving textures with either no spatial correlations [16] or experiments involving black-and-white textures [13, 18], and a nearly non-overlapping set of subjects.

We next examined the extent to which the multichannel nature of the model is critical to achieve a good correspondence to the experimental observations. We considered a simplified, one-channel model in which the first-stage mechanism was a threshold, sending darker-than-mean checks to 0 and lighter-than-mean checks to 1, and randomly assigned mid-gray checks. We also considered multichannel models in which one or more of the Silva-Chubb mechanisms were deleted. For the latter models, we either retained both F_1 and F_2, or deleted both – as their complementary, linearly-dependent nature was an important feature of the Silva-Chubb analysis [16]. In all cases, the model's second stage was adjusted to ensure that it produced thresholds for black-and-white textures that corresponded to previous psychophysical measurements [13, 18].

The last five rows of Fig. 8 show that overall, the predictions of these alternate models differ substantially from the measured thresholds. While the alternate models make similar predictions for $\vec{\sigma}_{(1\ \ 2)}$, and $\vec{\sigma}_{\left(\begin{smallmatrix}1 & \\ & 2\end{smallmatrix}\right)}$ (third and fifth columns), their predictions for the first-order statistic and the other second-order statistics differ widely from the psychophysical measurements. The model with binarization at mid-gray (Fig. 8C) and the model with luminance-like mechanism pair F_1 and F_2 (Fig. 8D) predict very high thresholds in some directions in the domains of $\vec{\sigma}_{(1)}$, $\vec{\sigma}_{(1\ \ 1)}$, and $\vec{\sigma}_{\left(\begin{smallmatrix}1 & \\ & 1\end{smallmatrix}\right)}$, in contrast to the psychophysical measurements and the predictions of the full model (Fig. 7). The other reduced models (Fig. 8E-G) do not predict unreasonably high thresholds for $\vec{\sigma}_{(1\ \ 1)}$ and $\vec{\sigma}_{\left(\begin{smallmatrix}1 & \\ & 1\end{smallmatrix}\right)}$, but nevertheless fail dramatically for $\vec{\sigma}_{(1)}$, and, in some cases (Fig. 8F,8G), for $\vec{\sigma}_{(1\ \ 1)}$ and $\vec{\sigma}_{\left(\begin{smallmatrix}1 & \\ & 1\end{smallmatrix}\right)}$ as well. Note that the rather strange predicted isodiscrimination contours for $\vec{\sigma}_{(1)}$ result not only from omitting mechanisms, but also from constraining the model's second stage to account for thresholds to spatially-correlated black-and-white textures.

Table 2. Quantification of Model Predictions

model	Experiment 1	Experiment 2	Experiment 3
full (F_1, F_2 F_3 F_4)	0.268	0.138	0.274
binarize at mid-gray	1.739	1.565	0.391
F_1 and F_2	0.713	0.573	2.840
F_1, F_2, and F_3	2.481	3.441	0.712
F_1, F_2, and F_4	1.769	0.231	0.856
F_3 and F_4	1.403	0.240	0.158
intersubject variability	0.585	0.130	0.212

Model predictions are quantified by the fraction of the variance of the threshold measurements accounted for by the model. For Experiment 1, only first- and second-order statistics are considered. The entry for intersubject variability is the median of the fraction of variance of one subject's data that is accounted for by a second subject. For further details, see Methods.

Table 2 (second column) quantifies the goodness of fit of the full model and the alternate models considered in Fig. 8, in terms of the fraction of variance unexplained by each model's prediction of the average psychophysical thresholds. The full model leaves 27% of the variance unexplained; the best alternate model (F_1 and F_2 only) leaves 71% of the variance unexplained. For the other models, more than 100% of the variance is unexplained (i.e., in terms of explained variance, the model is worse than a model that simply predicts that all thresholds are zero). The last row of Table 2 compares these statistics with a measure of intersubject variability, the median of the fraction of the variance unexplained in one subject's data, based on a second subject (see Methods). For the full model, the fraction of variance unexplained is comparable to the intersubject variability; for all the alternate models, the fraction of variance unexplained is greater than the intersubject variability, often substantially so.

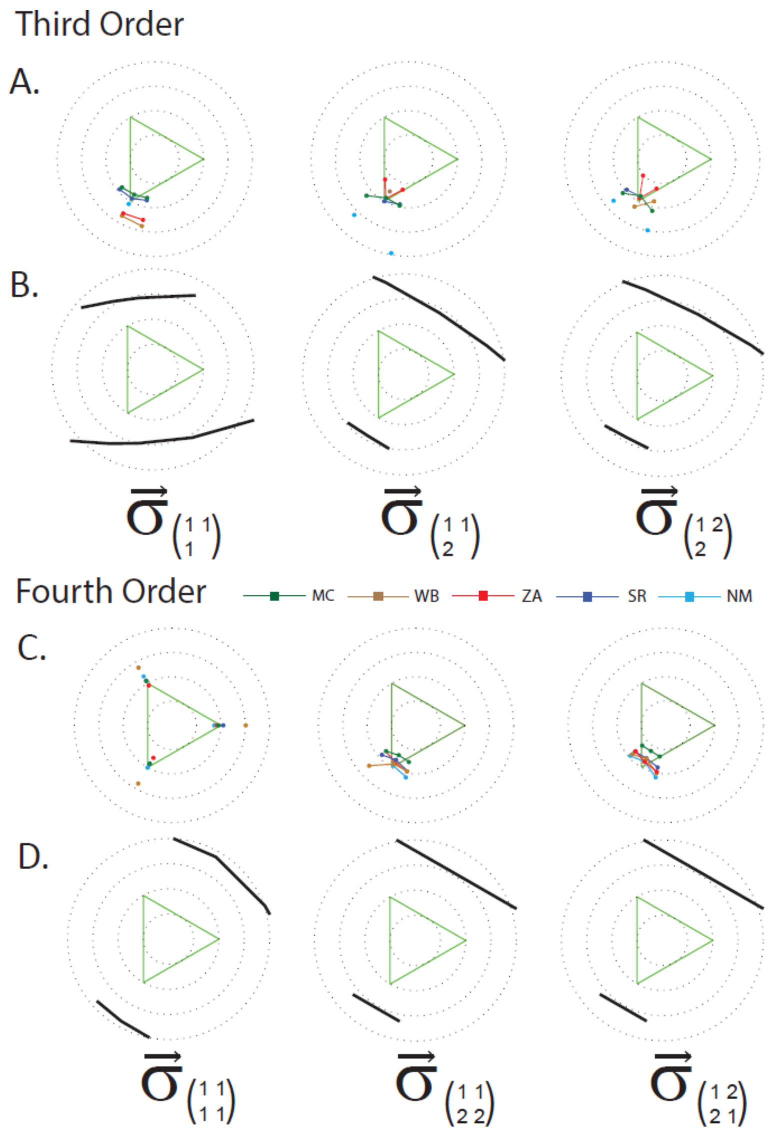

Fig. 9. Psychophysical thresholds and model predictions for Experiment 1 in the triangular domains of selected third- and fourth-order statistics. First and third rows show experimental measurements; second and fourth rows show model predictions. Points are disconnected if intervening directions correspond to chance performance (psychophysical data) or thresholds > 4 (model). For examples of the domains, see Fig. 3: the third-order domain $\vec{\sigma}_{\left(\begin{smallmatrix}1&1\\1&\end{smallmatrix}\right)}$ is shown in Fig. 3A; the fourth-order domain $\vec{\sigma}_{\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)}$ is shown in Fig. 3B. Other plotting conventions as in Fig. 8.

Thresholds for third- and fourth-order statistics are shown in the first and third rows of Fig. 9. Thresholds were generally much higher than for first- and second-order correlations, and there were many directions in the third- and fourth-order domains in which performance was at chance, even for maximally-correlated textures. The model predicts these higher thresholds, and largely accounts for the directions in which thresholds could be measured. For the third-order domains and the fourth-order domains $\vec{\sigma}_{\begin{pmatrix}1&1\\2&2\end{pmatrix}}$ and $\vec{\sigma}_{\begin{pmatrix}1&2\\2&1\end{pmatrix}}$, predicted thresholds are closest to the borders of the stimulus domain in the direction of the lower-left vertex, corresponding to configurations in which luminance values sum to zero (mod 3). These are the directions in which subjects' performance was better than chance. However, for the fourth-order domain $\vec{\sigma}_{\begin{pmatrix}1&1\\1&1\end{pmatrix}}$, subjects' performance was better than chance at each vertex, and this does not appear to be accounted for by the model.

4.2 Experiment 2

Experiment 2 examines how different kinds of image statistics combine, focusing on second-order image statistics. Studies were organized into four groups: group I made use of stimuli that combined $\vec{\sigma}_{(1\ \ 1)}$ and $\vec{\sigma}_{(1\ \ 2)}$; the other groups made use of stimuli that combined a horizontal correlation $\vec{\sigma}_{(1\ \ s)}$ with a vertical correlation $\vec{\sigma}_{\begin{pmatrix}1\\s'\end{pmatrix}}$. Each of these genera of correlations includes three species (corresponding to the three vertices of its domain), so pairwise combinations of genera encompassed multiple pairwise combinations of specific image statistics (Table S1).

Experimental results are shown in alternate rows in Fig. 10. All combinations of image statistics supported image segmentation, and the pattern of threshold behavior is consistent across subjects, as in Experiment 1 (Fig. 8A). Moreover, thresholds are again nearly symmetric for positive and negative correlation strengths, and isodiscrimination contours were very nearly circular or elliptical. These behaviors were also captured by the model of Fig. 7, including the orientation of the isodiscrimination contours in most cases.

Quantitatively (Table 2), the model of Fig. 7 also performs well; the fraction of variance unexplained (14%) is comparable to median intersubject variability (13%). None of the alternate models perform comparably: two (the models with mechanisms F_1, F_2, and F_4, and the model with mechanisms F_3 and F_4) have reasonable performance (23-24% of variance unexplained); the others perform very poorly.

4.3 Experiment 3

Experiment 3 determines sensitivity to a salient subset of second-order statistics across a range of gray levels. Specifically, we examined two kinds of correlations between adjacent checks: correlations that produced gradients (Fig. 5A), and correlations that produced streaks (Fig. 5B). Note that for a given number of gray levels, the luminance distribution is the same for the two kinds of stimuli, but the spatial organization differs.

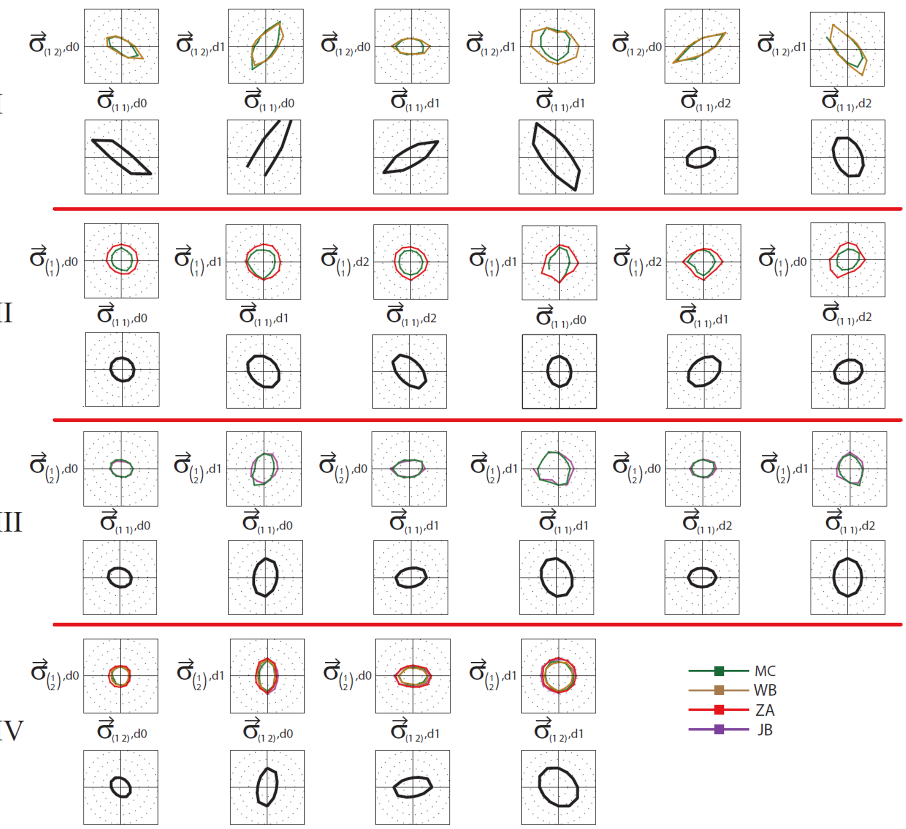

Fig. 10. Psychophysical thresholds and model predictions for Experiment 2, which combines pairs of second-order statistics. Each pair of rows corresponds to an experimental group, delineated in Table S1. The upper row of each group shows experimental measurements; the lower row shows model predictions. The domains in each group combine a second-order statistic drawn from two genera, shown at the beginning of the rows, as delineated in Table S1. Individual domains are labeled according to the correlation species varied along the ordinate and abscissa, each of which is defined by a vertex within the genus' triangular domain (e.g., Fig. 2). Barycentric coordinates of the vertices are indicated by d0 for $(1,0,0)$, d1 for $(0,1,0)$, and d2 for $(0,0,1)$. The first and fifth domains of Group I are shown in Fig. 4AB; the first domain and third domains of Group IV are shown in Fig. 4CD. Other plotting conventions as in Figs. 8 and 9.

Thresholds were measured in six subjects, using the same procedures as Experiments 1 and 2. One subject, MC (an author) was also a participant in Experiments 1 and 2.

For the gradients (Fig. 11A, left), thresholds were an inverted-U function of the number of gray levels, with texture-contrast thresholds of approximately $c = 0.5$ for 3 and 11 gray levels, and a maximal threshold of 0.7-1.0 for 5 gray levels. For the streaks (Fig. 11A, right), thresholds were 0.2-0.3 for all subjects, and independent of gray level. Thresholds were independent of the direction of the gradient or the orientation of the streak.

The thresholds predicted by the model are shown in Fig. 11B. The model predicts the inverted-U shape of the sensitivity function and its peak position, as well as the absolute thresholds at the extremes of the curve, but overestimates the threshold for 5 gray levels. For streaks, the model predicts the independence of gray levels and also predicts the absolute thresholds.

Fig. 11C shows the predictions of the alternate models in Fig. 8. With regard to the gradients, several of these models predict unreasonably high thresholds for 3 or 4 gray levels, far exceeding the psychophysical findings. The one model that does not fail in this fashion (F_3 and F_4 only) predicts a higher threshold for 7 gray levels rather than 5; all subjects

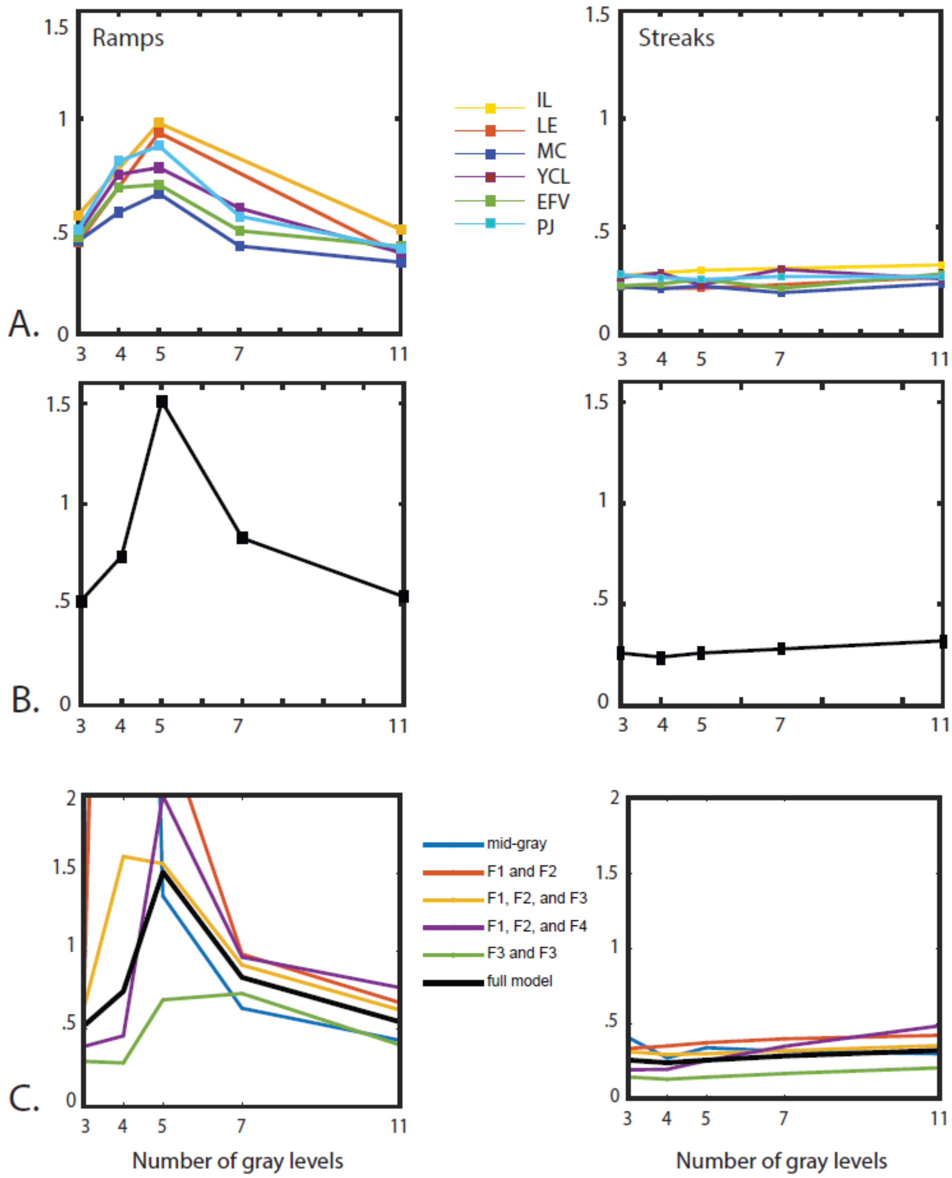

Fig. 11. Psychophysical thresholds and model predictions for Experiment 3, gradients (left) and streaks (right). A: Individual subjects' data. Uncertainties (2SEM) for individual subject thresholds are <0.05 and are not shown. B: Predictions of the model. C: Predictions (color) of alternate models whose first stage consists of a single mechanism with a threshold at mid-gray, or a subset of the Silva-Chubb mechanisms. The predictions of the full model are shown in black. Off-scale points correspond to predicted thresholds > 4.

showed the opposite behavior. With regard to streaks, most of the alternate models correctly predicted the finding that thresholds were approximately independent of the number of gray levels. However, the threshold predicted by the F_3 and F_4 only-model was lower than measured, and the thresholds predicted by the F_1, F_2, and F_4 model showed a near-doubling of threshold as the number of gray levels increased from 3 to 11, also inconsistent with the data.

In a quantitative analysis (Table 2), the accuracy of model predictions (27% of variance unexplained) is comparable to intersubject variability (21% of variance unexplained). Most of the alternate models do not perform as well. The model with mechanisms F_3 and F_4 provides a better fit in this experiment (16% of the variance unexplained), but, as noted above, this model performs very poorly in Experiment 1 (none of the variance explained).

5. Discussion

This study examined sensitivity to visual textures, with the broad goal of understanding how neural computations analyze a high-dimensional sensory space. Visual textures constitute such a space, as their parametric description – local image statistics – includes not only the luminance distribution, but also spatial correlations among pairs of checks, triplets, etc. While human visual sensitivity to image statistics is selective [24], the number of perceptually relevant texture parameters is quite large. The visual system can detect texture differences based on several aspects of the luminance distribution in the absence of spatial correlations, and multiple kinds of spatial correlations, even for textures that only have black and white elements. Moreover, typical textures have spatial correlations *and* are not restricted to two luminance levels; the number of parameters required to describe spatial correlations grows rapidly as the number of luminance levels increase [20]. Here, to understand how such textures are processed, we examine perceptual thresholds for discrimination of several classes of textures that have multiple gray levels and also spatial correlations.

Our main findings in Experiments 1 and 2 are that two observations previously made for black-and-white spatially-correlated textures [13] apply in this more general context: thresholds for negative and positive correlations are nearly equal, and signals from different local image statistics combine quadratically. The net result is that isodiscrimination thresholds, parameterized by local image statistics, are approximately elliptical.

Our main finding in Experiment 3 is that gray-level distribution and spatial correlations interact: for one kind of spatial correlation (gradients), threshold for discrimination from randomness has a sharp maximum when 5 gray levels are present; for a second kind of spatial correlation (streaks), the threshold is approximately independent of gray levels.

While this interaction is perhaps unsurprising, it provides empirical evidence that gray-level distributions and spatial correlations are not merely processed independently. In an attempt to capture how these dimensions interact, we constructed a computational model to account for our findings, based on previous studies of gray-level textures without spatial correlations, and studies in our lab of spatially-correlated black-and-white textures. As the model is fully constrained by those previous studies, it has no free parameters. The model reproduces the qualitative features of our findings --the elliptical shape of the isodiscrimination contours seen in Experiments 1 and 2, and the interaction between the number of gray levels and spatial correlations seen in Experiment 3 – and provides a reasonable quantitative prediction of the thresholds as well.

We note, however, that all of the stimuli are constructed with discrete, monochrome checks, so it is an open issue as to whether the approach extends to textures with continuous gradations and/or chromatic content.

5.1 The model

The proposed model (Fig. 7) combines elements that process luminance distributions and elements that process spatial pattern in a novel manner. We emphasize that it is a computational model; its components are not intended to have direct physiologic correlates.

The first stage of the model consists of a set of parallel mechanisms that process luminance distributions. Each mechanism transforms the visual input into a binary representation, assigning each check to 0 or 1 with a probability determined by the gray-level of the stimulus. The characteristics of these mechanisms – i.e., the way that the probability depends on gray levels (F_m , eq. (6))and their relative strengths (w_m , eq. (17))-- are taken from the studies of Silva and Chubb [16], in which they used a search task to measure discrimination of spatially-uncorrelated gray-level textures. By using the Silva and Chubb mechanisms, our model is guaranteed to reproduce the key findings of Chubb and colleagues [14, 16, 17, 33]: that spatially-uncorrelated gray-scale textures form a three-dimensional perceptual space, and textures that are indistinguishable by these mechanisms are perceptually indistinguishable.

The second stage confers the sensitivity to spatial correlations. Each of the internal representations produced by the first stage is analyzed by mechanisms sensitive to patterns in 2×2 clusters of checks. The second stage is thus sensitive not only to pairwise correlations, but also to third- and fourth-order correlations among nearest neighbors, as is needed to account for early observations concerning isodipole textures [34]. The way that signals from these local correlations combine (eq. (17)) is fully constrained by the requirement that the model accounts for discrimination thresholds for black-and-white textures, previously measured in our lab [13, 18, 27, 28], as detailed in the Supplemental Document, "Specification of the model's quadratic form."

Our model can be viewed as a generalization of a "back-pocket" [15] framework: its first stage consists of several independent analyzers and their outputs are combined quadratically. But in contrast to the standard back-pocket model, the outputs of the analyzers are multivariate quantities that contain spatial information, rather than scalars. Correspondingly, the quadratic combination rule is a quadratic form, rather than a simple square law. This generalization allows for interactions between gray-level distributions and spatial pattern.

The ability of the model to predict our findings, both qualitatively and quantitatively, depends not only on its overall structure, but also on the specifics of the model's first stage: the four independently-identified Silva-Chubb mechanisms [16]. When these mechanisms are replaced by a simple threshold, or, when one or more of them are removed, the elliptical isodiscrimination contours of Experiments 1 and 2 are lost, and some thresholds are predicted to be unreasonably large (Fig. 8C-G). These alternate models also are not able to account for the interaction of the gray-level distribution and spatial correlations seen in Experiment 3 (Fig. 11). Alternate models also fall short in terms of quantitative prediction of the measured thresholds for first- and second-order statistics and their combinations (Table 2).

5.2 Simplifications and approximations

In keeping with the goal of focusing on the structure of the computations underlying texture processing and avoiding an explosion of parameters, the model makes substantial simplifications regarding the neural circuitry underlying spatial processing. Center-surround organization and orientation-tuned spatial filtering are not explicitly modeled. Instead, the net effect of checks surrounding the central element are lumped together into the stochastic threshold that converts the central check into an internal binary representation. As we don't model "receptive fields" explicitly, we don't take into account eccentricity-dependence of receptive field centers and surrounds (and consequent eccentricity-dependent changes in the typical number of checks within a receptive field). Finally, the nearest-neighbor correlations that define the textures necessarily induce longer-range correlations, but these are neglected –

the model's sensitivity to spatial structure is determined only by configurations in a 2×2 block of checks, independent of eccentricity.

These simplifications enable us to constrain the model based on previous studies – though this too entails some assumptions. The first-stage mechanisms are taken from previous studies with spatially-uncorrelated textures [16]. This makes the assumption that these mechanisms are unchanged when spatial correlations are present, and when the specific gray level distributions differ substantially. Further, the second-stage mechanisms we use to model the processing of spatial structure were determined from studies in which structured textures were discriminated from random ones [13]. Here they are applied to internal representations in which the comparison is between two non-random textures. The "translation invariance" needed for this generalization (i.e., that discrimination between textures with coordinates $\vec{y}$ and $\vec{z}$ depends only on $\vec{y}-\vec{z}$) is only approximate [28]. Moreover, these internal representations, though binary, are outside the stimulus set used in [13]: because of the action of the stochastic threshold, they are no longer maximum-entropy.

Despite these approximations and simplifications, the agreement of the model with the experimental data is good -- but there are also specific systematic discrepancies that are larger than intersubject variability. Overall, the model underestimates the thresholds for first- and some second-order statistics, and overestimates the threshold for third- and fourth-order statistics. For some first- and second-order statistics, the orientation of the isodiscrimination ellipse is also not accurately predicted.

While any of the above approximations and simplifications may contribute to the model's inaccuracies, the overall under-prediction of low- order thresholds and over-prediction of high-order thresholds is expected to be very sensitive to the precise shapes of the operating curves of the first-stage mechanisms. Specifically, if the thresholds were slightly less stochastic – i.e., the curves transitioned more rapidly from 0 to 1 – then the balance would tilt towards the high-order correlations, as these rely on preservation of the image structure in multiple neighboring checks. The precise shapes of the first-stage mechanisms will also influence the orientation of the ellipses, as well the replacement of true surround subtraction by a stochastic threshold, as well as the neglect of correlations at larger spatial scales.

While our focus is on a simple conceptual model for visual computations, the model's structure is fully compatible with more elaborate, physiologically-realistic models. Our main building blocks – linear summation and pointwise nonlinearities – are typical building blocks of such models. As mentioned above, the stochastic threshold is an approximation of the influence of the receptive field surround. Moreover, the nonlinearities required to extract third- and fourth-order spatial correlations are known to exist in primate area V2 [35], and emerge naturally in models of recurrent neural networks[36].

5.3 How visual modalities combine

An important way in which the brain copes with the complexity of the visual world is to utilize separate regions or networks specialized for processing of visual modalities, such as orientation, color, shape, motion, and depth. While initial studies emphasized specialization and modularity[2, 37, 38], it is now well-recognized that these modules are not independent, as subsequent studies revealed both physiological and psychophysical evidence for intermixing [8, 9, 39-41].

The model structure we propose presents a common theme for the way in which cross-modal interactions are structured. In our model, local luminance is processed by a parallel set of mechanisms, each of which provides an internal representation that is then analyzed by a second stage, which is sensitive to spatial correlation. Similarly, in Papathomas' [39] study of chromatic interactions with motion, local chromatic signals provide tags, which is then used by a standard spatiotemporal analyzer to extract unambiguous motion. Non-Fourier motion can be viewed in the same way: it can be detected by a cascade in which local flicker or local unsigned contrast becomes tokens that serve as a starting point for standard motion analysis[42]. Finally, in studies of structure-from-motion [43], the spatial arrangement of locally-extracted motion signals constitute an internal representation that is then analyzed for shape.

Our model is a further elaboration on this theme. In the first stage of our model, multiple internal representations are abstracted from the luminance image. At the model's second stage, each of these internal representations undergoes a spatial analysis. Each of these transformations is both local and nonlinear, but the nonlinearities address different aspects of the input: luminance distribution and spatial structure. The net result is a computation that could not be achieved by independent processing within these modalities.

5.4 Relevance to visual processing of natural scenes

The computation captured by this model is central to efficient visual processing of natural scenes. As has been proposed by the efficient coding hypothesis [44], the visual system is tuned to take advantage of the distinctive statistical characteristics of natural visual inputs. These characteristics include not only their well-known $1/f^2$ spatial power spectra [45-47], but also, their luminance and local image statistics [23-25, 48]. Specifically, some kinds of local image statistics are quite variable across natural scenes, and are therefore highly informative, while others are relatively more stereotyped and/or predictable, and therefore less informative. Importantly (and perhaps surprisingly), these previous studies [23-25] have shown that the informativeness of different kinds of local image statistics in natural scenes is closely correlated with visual sensitivity to these statistics when they are isolated in our synthetic textures. Our model shows that the computations that implement efficient coding can be accomplished in a compact fashion, that is, by combining the outputs of a small number of local mechanisms (the first stage of the model) with a single quadratic nonlinearity (the second stage of the model).

Note also that our findings are inconsistent with the notion that visual sensitivity to an image statistic merely reflects the extent to which the statistic reduces entropy. All texture domains have a fully random (maximally entropic) texture at the origin, and, for small texture contrasts, the reduction in entropy depends only on the distance from the origin (Appendix B of [18]), independent of the domain or the direction of the displacement. The widely varying sensitivities we observe, and the elliptical rather than circular isodiscrimination contours, indicate that sensitivity varies widely across image statistics, even though they each reduce entropy by the same amount. This selectivity is inconsistent with coding entropy reduction *per se*, but, as mentioned above, corresponds instead to the efficient coding of natural scenes.

Finally, we note that the efficient coding framework is also relevant to understanding how the chromatic content of natural scenes is processed [49-52]; however, the present analysis (and that of many others [21, 53-56]) is limited to their achromatic aspects.

6. Back matter

Funding. National Institutes of Health NEI EY07977

Acknowledgments. Portions of this work were presented at the 2017 and 2018 meetings of the Vision Sciences Society (St. Petersburg, FL) and the 2017 and 2018 meetings of the Society for Neuroscience (Washington, DC and San Diego, CA). A portion of the psychophysical data presented in Experiments 1 and 2 has also been presented in [21]. We thank Lilah Evans for assistance with data collection for Experiment 3. This work was supported by NIH NEI EY7977.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplemental Document for supporting content.

7. References

1. H. S. M. Coxeter, *Introduction to Geometry Second Edition* (John Wiley & Sns, 1969).
2. D. Van Essen, and J. Maunsell, "Hierarchical organization and functional streams in the visual cortex," Trends in Neuroscience **6**, 370-375 (1983).
3. M. Livingstone, and D. Hubel, "Segregation of form, color, movement, and depth: anatomy, physiology, and perception," Science **240**, 740-749 (1988).
4. D. H. Hubel, and M. S. Livingstone, "Segregation of form, color, and stereopsis in primate area 18," J Neurosci **7**, 3378-3415 (1987).
5. E. A. DeYoe, and D. C. Van Essen, "Concurrent processing streams in monkey visual cortex," Trends Neurosci **11**, 219-226. (1988).
6. T. Yoshioka, and B. M. Dow, "Color, orientation and cytochrome oxidase reactivity in areas V1, V2 and V4 of macaque monkey visual cortex," Behav Brain Res **76**, 71-88. (1996).
7. T. A. Nealey, and J. H. Maunsell, "Magnocellular and parvocellular contributions to the responses of neurons in macaque striate cortex," J Neurosci **14**, 2069-2079 (1994).
8. W. H. Merigan, and J. H. Maunsell, "How parallel are the primate visual pathways?," Annu Rev Neurosci **16**, 369-402 (1993).
9. A. G. Leventhal, K. G. Thompson, D. Liu, Y. Zhou, and S. J. Ault, "Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex," J Neurosci **15**, 1808-1818. (1995).
10. E. N. Johnson, M. J. Hawken, and R. Shapley, "Cone inputs in macaque primary visual cortex," J Neurophysiol **91**, 2501-2514 (2004).
11. T. Maddess, Y. Nagai, J. D. Victor, and R. R. Taylor, "Multilevel isotrigon textures," J Opt Soc Am A Opt Image Sci Vis **24**, 278-293 (2007).
12. T. Maddess, and Y. Nagai, "Discriminating isotrigon textures," Vision Res **41**, 3837-3860 (2001).
13. J. D. Victor, D. J. Thengone, S. M. Rizvi, and M. M. Conte, "A perceptual space of local image statistics," Vision Res **117**, 117-135 (2015).
14. C. Chubb, J. Econopouly, and M. S. Landy, "Histogram contrast analysis and the visual segregation of IID textures," J Opt Soc Am A Opt Image Sci Vis **11**, 2350-2374 (1994).
15. C. Chubb, and M. Landy, "Orthogonal distribution analysis: a new approach to the study of texture perception," in *Computational models of visual processing*, M. S. Landy, and Movshon, J.A., ed. (MIT Press, 1991), pp. 291-301.
16. A. E. Silva, and C. Chubb, "The 3-dimensional, 4-channel model of human visual sensitivity to grayscale scrambles," Vision Res **101**, 94-107 (2014).
17. C. Chubb, M. S. Landy, and J. Econopouly, "A visual mechanism tuned to black," Vision Res **44**, 3223-3232 (2004).
18. J. D. Victor, and M. M. Conte, "Local image statistics: maximum-entropy constructions and perceptual salience," J Opt Soc Am A Opt Image Sci Vis **29**, 1313-1345 (2012).
19. J. D. Victor, M. M. Conte, and C. F. Chubb, "Textures as Probes of Visual Processing," Annual review of vision science **3**, 275-296 (2017).
20. R. M. Haralick, Shanmugam, K., Dinstein, I., "Textural features for image classification," IEEE Trans. Systems, Man and Cybernetics. **SMC-3**, 610-621 (1973).
21. J. Portilla, and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," International Journal of Computer Vision **40**, 49-71 (2000).

22. S. C. Zhu, Y. Wu, and D. Mumford, "Filters, random fields and maximum entropy (FRAME): Towards a unified theory for texture modeling," International Journal of Computer Vision **27**, 107-126 (1998).
23. T. Tesileanu, M. M. Conte, J. J. Briguglio, A. M. Hermundstad, J. D. Victor, and V. Balasubramanian, "Efficient coding of natural scene statistics predicts discrimination thresholds for grayscale textures," eLife **9** (2020).
24. G. Tkacik, J. S. Prentice, J. D. Victor, and V. Balasubramanian, "Local statistics in natural scenes predict the saliency of synthetic textures," Proc Natl Acad Sci U S A **107**, 18149-18154 (2010).
25. A. M. Hermundstad, J. J. Briguglio, M. M. Conte, J. D. Victor, V. Balasubramanian, and G. Tkacik, "Variance predicts salience in central sensory processing," eLife **3** (2014).
26. J. D. Victor, C. Chubb, and M. M. Conte, "Interaction of luminance and higher-order statistics in texture discrimination," Vision Res **45**, 311-328 (2005).
27. J. D. Victor, D. J. Thengone, and M. M. Conte, "Perception of second- and third-order orientation signals and their interactions," J Vis **13**, 21 (2013).
28. J. D. Victor, S. M. Rizvi, and M. M. Conte, "Two representations of a high-dimensional perceptual space," Vision Res **137**, 1-23 (2017).
29. B. Julesz, and J. R. Bergen, " Textons, the fundamental elements in preattentive vision and perception of textures," The Bell System Technical Journal **62**, 1619-1645 (1983).
30. B. Julesz, "Textons, the elements of texture perception, and their interactions," Nature **290**, 91-97 (1981).
31. S. S. Wolfson, and M. S. Landy, "Examining edge- and region-based texture analysis mechanisms.," Vision Res **38**, 439-446 (1998).
32. L. Frisen, and A. Glansholm, "Optical and neural resolution in peripheral vision," Invest Ophthalmol **14**, 528-536 (1975).
33. C. Chubb, J. H. Nam, D. R. Bindman, and G. Sperling, "The three dimensions of human visual sensitivity to first-order contrast statistics," Vision Res **47**, 2237-2248 (2007).
34. B. Julesz, E. N. Gilbert, and J. D. Victor, "Visual discrimination of textures with identical third-order statistics," Biol Cybern **31**, 137-140 (1978).
35. Y. Yu, A. M. Schmid, and J. D. Victor, "Visual processing of informative multipoint correlations arises primarily in V2," eLife **4** (2015).
36. J. Joukes, Y. Yu, J. D. Victor, and B. Krekelberg, "Recurrent Network Dynamics; a Link between Form and Motion," Front Syst Neurosci **11**, 12 (2017).
37. S. Shipp, and S. Zeki, "Segregation of pathways leading from area V2 to areas V4 and V5 of macaque monkey visual cortex," Nature **315**, 322-325 (1985).
38. M. S. Livingstone, and D. H. Hubel, "Psychophysical evidence for separate channels for the perception of form, color, movement, and depth," J Neurosci **7**, 3416-3468 (1987).
39. T. V. Papathomas, A. Gorea, and B. Julesz, "Two carriers for motion perception: color and luminance," Vision Res **31**, 1883-1892 (1991).
40. K. R. Gegenfurtner, and M. J. Hawken, "Interaction of motion and color in the visual pathways," Trends Neurosci **19**, 394-401 (1996).
41. I. Rentzeperis, A. R. Nikolaev, D. C. Kiper, and C. van Leeuwen, "Distributed processing of color and form in the visual cortex," Frontiers in psychology **5**, 932 (2014).
42. C. Chubb, and G. Sperling, "Processing stages in non-Fourier motion perception," Invest Ophth Vis Sci **29**, 266 (1988).
43. D. Regan, D. Giaschi, J. A. Sharpe, and X. H. Hong, "Visual processing of motion-defined form: selective failure in patients with parietotemporal lesions," J Neurosci **12**, 2198-2210 (1992).
44. H. B. Barlow, ed. *Possible principles underlying the transformation of sensory messages* (MIT Press, 1961).
45. D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," J Opt Soc Am [A] **4**, 2379-2394 (1987).

46. X. Kuang, M. Poletti, J. D. Victor, and M. Rucci, "Temporal encoding of spatial information during active visual fixation," Curr Biol **22**, 510-514 (2012).

47. Y. Dan, J. J. Atick, and R. C. Reid, "Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory," J Neurosci **16**, 3351-3362 (1996).

48. V. Balasubramanian, and P. Sterling, "Receptive fields and functional architecture in the retina," J Physiol **587**, 2753-2767 (2009).

49. S. M. Nascimento, K. Amano, and D. H. Foster, "Spatial distributions of local illumination color in natural scenes," Vision Res **120**, 39-44 (2016).

50. C. A. Parraga, G. Brelstaff, T. Troscianko, and I. R. Moorehead, "Color and luminance information in natural scenes," J Opt Soc Am A Opt Image Sci Vis **15**, 563-569 (1998).

51. G. J. Burton, and I. R. Moorhead, "Color and spatial structure in natural scenes," Appl Opt **26**, 157-170 (1987).

52. G. Buchsbaum, and A. Gottschalk, "Trichromacy, opponent colours coding and optimum colour information transmission in the retina," Proc R Soc Lond B Biol Sci **220**, 89-113 (1983).

53. J. Freeman, C. M. Ziemba, D. J. Heeger, E. P. Simoncelli, and J. A. Movshon, "A functional and perceptual signature of the second visual area in primates," Nat Neurosci **16**, 974-981 (2013).

54. M. Carandini, J. B. Demb, V. Mante, D. J. Tolhurst, Y. Dan, B. A. Olshausen, J. L. Gallant, and N. C. Rust, "Do we know what the early visual system does?," J Neurosci **25**, 10577-10597 (2005).

55. B. A. Olshausen, and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," Nature **381**, 607-609. (1996).

56. E. P. Simoncelli, and B. A. Olshausen, "Natural image statistics and neural representation," Annu Rev Neurosci **24**, 1193-1216 (2001).