Vision Research 158 (2019) 208–220

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

The features that control discrimination of an isodipole texture pair

Kier Groulx[a], Charles Chubb[a,*], Jonathan D. Victor[b], Mary M. Conte[b]

[a] Department of Cognitive Sciences, University of California at Irvine, United States
[b] Weill Cornell Medical College, United States

ARTICLE INFO

No. of Reviewers – 2

Keywords:
Isodipole texture
Centroid method
Preattentive vision

ABSTRACT

Visual features such as edges and corners are carried by high-order statistics. Previous analysis of discrimination of isodipole textures, which isolate specific high-order statistics, demonstrates visual sensitivity to these statistics but stops short of analyzing the underlying computations. Here we use a new texture centroid paradigm to probe these computations. We focus on two canonical isodipole textures, the even and odd textures: any 2×2 block of even (odd) texture contains an even (odd) number of black (and white) checks. Each stimulus comprised a spatially random array of black-and-white texture-disks (background = mean gray) that varied in their fourth-order statistics. In the Even (Odd) condition, disks varied along the continuum between random coinflip texture and pure (highly structured) even (odd) target texture. The task was to mouse-click the centroid of the disk array, weighting each disk location by the target structure level of the disk-texture (ranging from 0 for coinflip to 1 for even or odd). For each of block-sizes $S = 2 \times 2$, 2×3, 2×4 and 3×3, a linear model was used to estimate the weight exerted on the subject's responses by the differently patterned blocks of size S. Only the results with 2×4 and 3×3 blocks were consistent with the data. In the Even condition, homogeneous blocks exerted the most weight; in the odd condition, block-pattern symmetry was important. These findings show that visual mechanisms sensitive to four-point correlations do not compute evenness or oddness *per se*, but rather are activated selectively by features whose frequency varies across isodipole textures.

1. Introduction

Extraction of basic visual features depends not only on differences in luminance at two points (spatial contrast), but also on sensitivity to patterns of luminance at three or more points, also known as "higher-order correlations" (Textons, 1981; Morrone, Burr, & Maffei, 1982; Oppenheim & Lim, 1981). Paradigms that utilize isodipole textures constitute a principled approach for probing this sensitivity, since isodipole textures isolate higher-order correlations: by definition, they are matched in terms of their two-point correlations but differ in terms of three- or four-point correlations. The goal of the present study is to determine the nature of the computations that underlie visual performance on tasks involving these textures. We focus on the "even" and "odd" family of isodipole textures (Julesz, Gilbert, & Victor, 1978), as these are a well-studied class of isodipole textures whose structure is defined by a four-point correlation among 2×2 blocks of checks, and whose salient visual features include edges and corners. An example of the even (odd) texture is shown in the rightmost (leftmost) panel of Fig. 1.

Despite extensive study (Victor & Conte, 1989, 1991, 2012; Victor, Chubb, & Conte, 2005, 2015), the neural computations that underlie visual performance driven by the even vs. odd distinction are as yet unknown. One possibility is that neural circuits extract the four-point correlation that defines these textures. This requires a multiplication of luminance values at four points. However, while primary and secondary visual cortices appear to be the locus of the relevant computations (Victor, 1986; Purpura, Victor, & Katz, 1994; Yu, Schmid, & Victor, 2015), four-point multiplication does not readily map to known properties of cortical neurons. A possible resolution of this discrepancy is that visual performance does not rest on extraction of a four-point correlation *per se*. Other computations could serve as well, provided that they extracted statistical features implied by this correlation - for example, features that extend over larger regions of space, local symmetries, or the entropy of the probability distribution induced on local pattern types (Barbosa, Bubna-Litic, & Maddess, 2013; Maddess & Nagai, 2001; Nagai, Taylor, Loh, & Maddess, 2009; Taylor, Maddess, & Nagai, 2008).

To probe the underlying computations at a functional level, we use an extension of the "centroid paradigm" (Drew, Chubb, & Sperling, 2010; Sun, Chubb, Wright, & Sperling, 2016), a technique recently developed to investigate the filters that subjects can achieve for extracting the spatial distribution of various sorts of image statistics. As

* Corresponding author.
E-mail address: cfchubb@uci.edu (C. Chubb).

https://doi.org/10.1016/j.visres.2019.03.006
Received 19 December 2018; Received in revised form 9 March 2019; Accepted 12 March 2019
Available online 26 March 2019

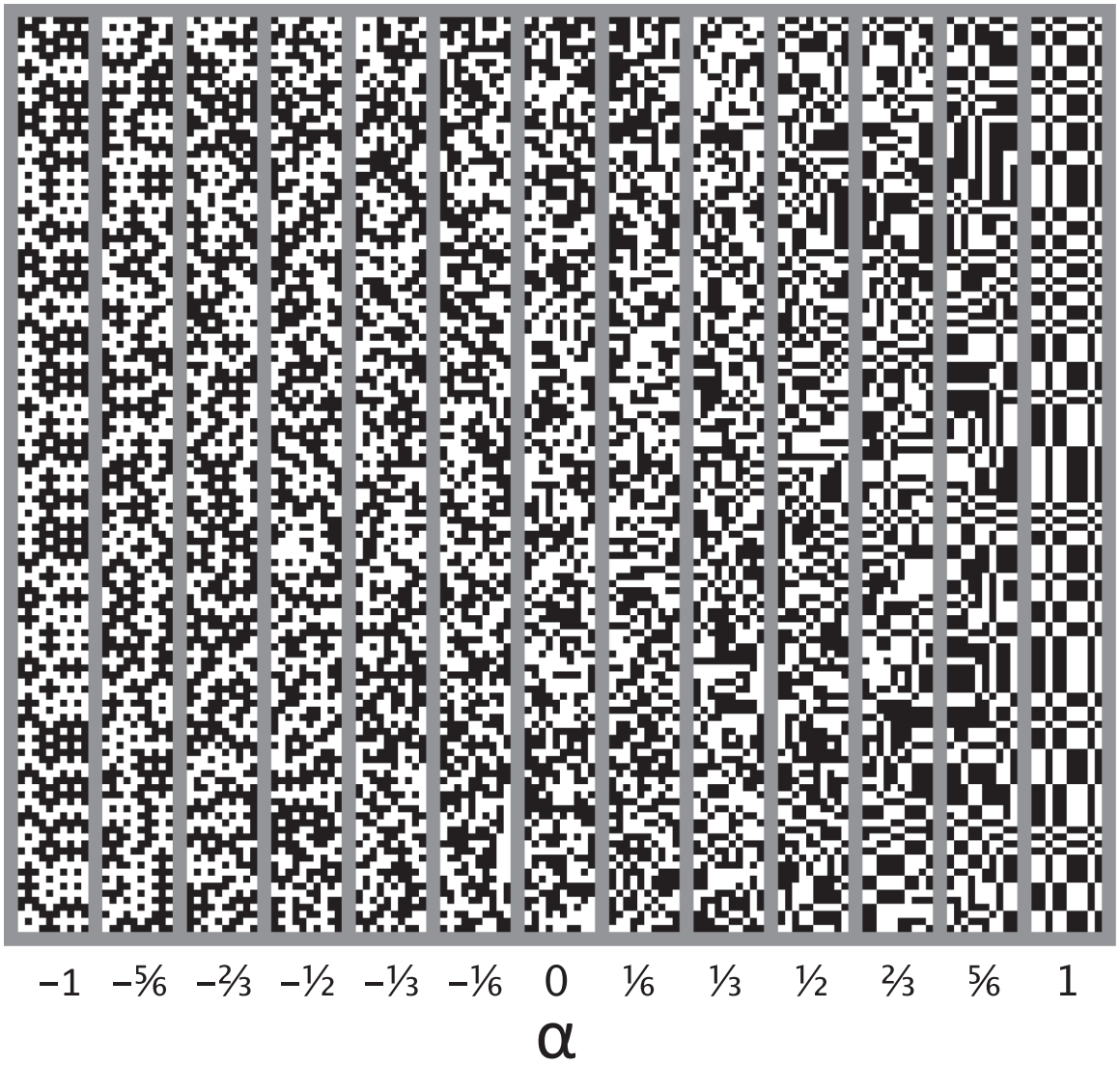

Fig. 1. Isodipole textures (Julesz et al., 1978). Different vertical bands of texture vary in the parameter α (Eq. (1)). The leftmost panel (with $\alpha = -1$) contains "odd" texture in which every 2×2 subblock contains an odd number of white (and black) checks. The rightmost panel contains the "even" texture in which every 2×2 subblock has an even number of white and black checks. More generally, the proportion of 2×2 subblocks with an even number of white and black checks in texture with a given level of α is $\frac{\alpha+1}{2}$.

we show, applying the centroid paradigm to stimuli consisting of patches of even and odd isodipole textures enables a characterization of the filter that the subject uses for selective attention. This characterization will indicate whether the visual system indeed computes the fourth-order correlation that defines the even vs. odd gamut, or rather, that the sensitivity to this distinction is best described in another way.

In a typical application of the centroid paradigm, stimuli consist of briefly flashed, spatially random scatterings of different sorts of items (e.g., dots of different colors or Gabor patches of different spatial frequencies and/or orientations). The subject is then tested in different "attention conditions," always with the same sorts of displays. In a given attention condition, the subject is instructed to mouse-click the centroid of the items in each display, giving weight to different types of items in accordance with a specified target filter. For example, in an experiment using dots of different colors, in one attention condition, the subject might be instructed to mouse-click the centroid of the red dots, ignoring (i.e., giving weight 0 to) dots of other colors; in another attention condition, the subject might be instructed to mouse-click the centroid of the blue dots, ignoring dots of other colors. The data from a given attention condition can be analyzed (using linear regression) to derive the function that gives the relative weight exerted on the subject's centroid estimates by all of the different types of items in the stimulus displays-what might be called the "filter" achieved by the subject in that condition. Since stimuli are identical in composition across different attention conditions, any differences in the observed filters across attention conditions must be due to differences in the attentional state of the subject.

In the present extension of the centroid paradigm, which we call the "texture centroid paradigm", the stimuli consist of disks that are filled with visual texture of the types shown in Fig. 2. Each disk in a given display contains a texture whose distribution lies along the gamut from a target distribution *T* and distractor distribution *D*. The task is to mouse-click the centroid of the display, weighing the location of each texture-disk in proportion to its signal strength, which is given by the distance of its distribution along the continuum from *D* to *T*. In order to perform well in a given condition of the texture centroid task, the subject must recruit the neurons in his or her visual system to produce a "filter" whose activation by different texture disks matches their signal strengths as closely as possible. The aim of the current study is to analyze the filters deployed by subjects in the context of the texture centroid task to differentiate (1) even texture (serving as *T*) from texture with $\alpha = 0$ (serving as *D*) and (2) odd texture (serving as *T*) from

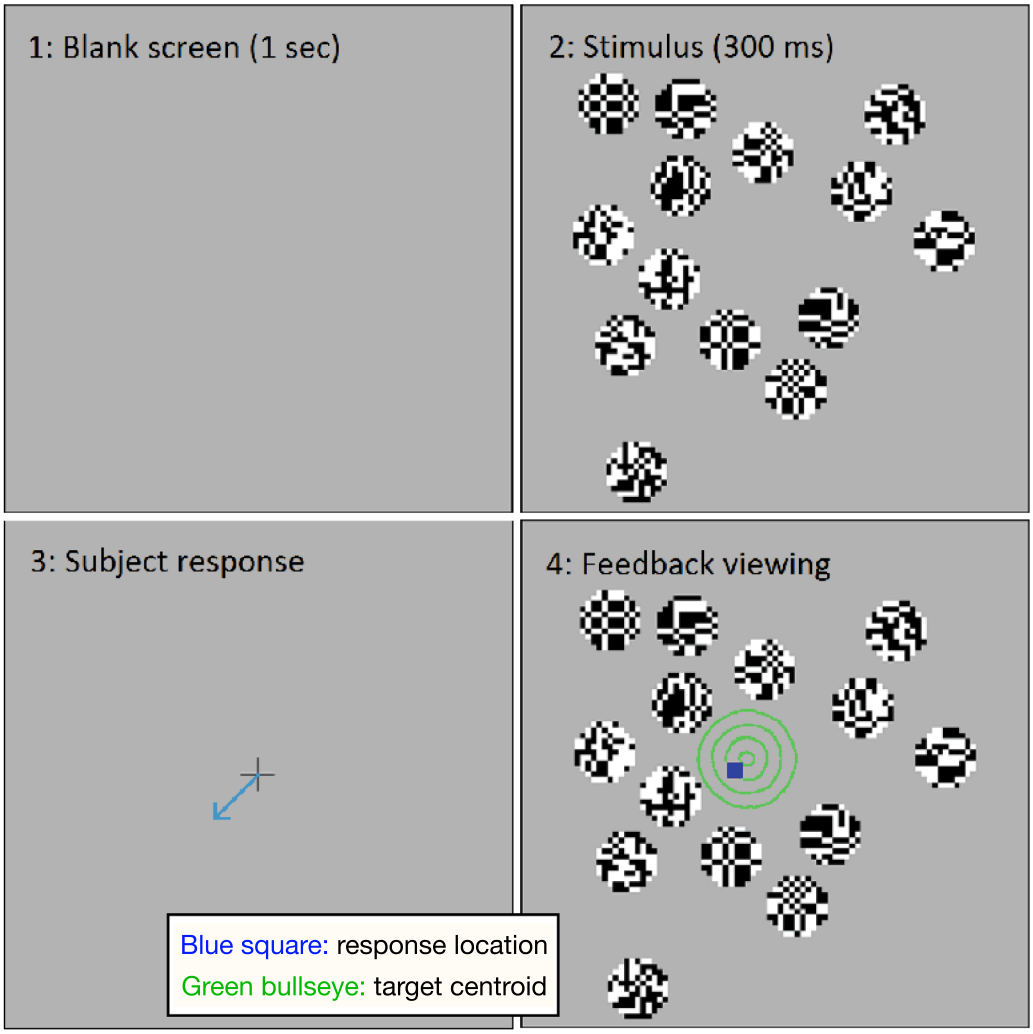

Fig. 2. The sequence of display images which occur during a trial of the centroid task. After a subject indicates that he or she is ready, the top left panel displays a mean gray field on the screen for one second. Afterwards, the top right stimulus panel displays a number of objects on the screen. The stimulus then disappears and a crosshair guided by the subject's mouse movements appears, allowing him or her to indicate the estimated centroid. Finally, the bottom right panel contains all objects from the original display, the subject's response, and a feedback bullseye centered on the correct response.

texture with $\alpha = 0$ (serving as D).

The disks used in our displays vary in a texture parameter called α (Victor et al., 2005; Victor & Conte, 2012). The "even" and "odd" textures are the opposite ends of a gamut parameterized by α, where $\alpha = 1$ corresponds to "even" and $\alpha = -1$ corresponds to "odd". A 2×2 block of abutting checks is said to have even (odd) parity if it contains an even (odd) number of black (or white) checks. For any texture disk d in any of the stimuli used in either of these two task conditions, we define

$$\alpha(d) = 2(p(d) - 0.5), \tag{1}$$

where $p(d)$ is the proportion of abutting 2×2 blocks of checks in d with even parity. Note that the extreme values of $\alpha(d)$, -1 and 1, correspond to probabilities $p(d)$ of 0 and 1, respectively, and $\alpha(d) = 0$ corresponds to $p(d) = 0.5$, the value for a random checkerboard.

In one task condition (the "Even" condition), the target weight of a texture disk d (i.e., the weight the subject strives to give d's location in the centroid computation) is equal to $\alpha(d)$, and in the other condition (the "Odd" condition), d's target weight is $-\alpha(d)$. Thus, in each task condition, the only image features that are explicitly relevant to performance are the 16 different possible patterns of 2×2 blocks of checks that can occur. In the Even condition, the ideal filter would give equal weight to all 2×2 block patterns with an even number of black checks and weight 0 to all 2×2 patterns with an odd number of black checks, and in the Odd condition, the ideal filter would give equal weight to all 2×2 patterns with an odd number of black checks and weight 0 to all 2×2 patterns with an even number of black checks.

As we shall see, in each condition, the filters achieved by our subjects deviate systematically from the ideal filter. These deviations provide important insights into the features in these textures to which human vision is actually sensitive.

2. Methods

Five subjects (two male) participated in this study at the University of California, Irvine. Subject $S1$ was the first author. The other subjects were inexperienced both in the centroid task and also in experiments using isodipole textures. The subjects participated under a UCI-approved IRB protocol with written consent.

Each stimulus consisted of a bounding box subtending 14×14 degrees of visual angle populated with 14 disks each of which subtended 1.06 degrees of visual angle. Each texture disk had an 11-check diameter and comprised 97 checks, each colored either black or white. That disks of this size are sufficiently large to analyze human sensitivity to variations in α is suggested by the results of Maddess and Nagai, 2001 who showed that discrimination of isodipole textures is invariant

down to targets of about 64 pixels. In addition, it should be noted that Victor and Conte (1989) found that the threshold for discriminating non-zero α from texture with $\alpha = 0$ is invariant with respect to check size for check-widths from 4 to 16 min, a range which includes the 5.8 min checks used here.

The x and y coordinates $C_x(d)$ and $C_y(d)$ of the disks d presented on a given trial were drawn from a circular, bivariate Gaussian distribution, but were subject to additional constraints that had the effect of making the distribution of disk-centers closer to uniform than to Gaussian. These additional constraints were:

1. For M_x and M_y are the mean x and y coordinates of the centers of all disks in the display,

$$\sqrt{\frac{1}{14}\sum_{\text{all disks } d}(C_x(d)-M_x)^2+(C_y(d)-M_y)^2} \approx 2.67^\circ \quad (2)$$

   In practice, this statistic varied from around 2.64 to 2.70° across different stimuli.
2. All disks were separated from each other by at least 3 check-widths.

In the Even condition, two (randomly selected) disks d in each display had $\alpha(d) = a$ for $a = 0, \frac{1}{6}, \cdots, 1$, and the subject strove (with trial-by-trial feedback) to mouse-click the centroid of the disks, giving weight to the location of each disk d proportional to $\alpha(d)$ (Eq. (1)). In the Odd condition, two disks d in each display had $\alpha(d) = s$ for $s = 0, -\frac{1}{6}, \cdots, -1$, and the subject strove to mouse-click the centroid of the disks, giving weight to the location of each disk d proportional to $-\alpha(d)$.

The level of α in a given disk was carefully controlled. Of the 97 checks contained in a given disk, 76 checks were the "lower-right anchors" of 2 × 2 blocks contained within the disk: i.e., for each of these checks h, the checks directly (a) above, (b) to the left and (c) diagonally above and to the left of h were also in the disk. For any given check h in this lower-right anchoring set, the parity of the 2 × 2 block anchored by h was determined ahead of time. Specifically, a vector of length 76 was filled with $n_{even} = \text{round}\left(76 \times \frac{\alpha-1}{2}\right)$ instances of the value 1 and $76 - n_{even}$ instances of -1. This vector was then randomly scrambled, and the values were assigned to the 76 lower-right anchors h by a function ParityController(h). The 21 checks not in the lower-right anchoring set were randomly and independently assigned the values equal to -1 or 1 with equal probability. Then the values of the lower-right anchoring checks h were assigned as follows:

$$\text{value of check } h = A \times B \times C \times \text{ParityController}(h) \quad (3)$$

where A, B and C are the previously determined values of the checks directly above, left of and above-left of h.

A sample of texture disks used in the Even task condition is shown in the top row of Fig. 3, and a sample of texture disks used in the Odd condition is shown in the bottom row.

Fig. 2 shows a trial in the Even condition. (Not shown is the "bounding box", a thin, black line that circumscribed the square region of the monitor screen in which the stimulus was displayed. The bounding box remained visible throughout the trial.) Subjects were given no special instructions about where to maintain fixation or where to deploy their attention, and no effort was made to monitor their eye movements during the experiment. The subject first viewed a blank (mean gray) field circumscribed by the bounding box for 1 s. The stimulus was then presented for 300 ms. Then there appeared a blank field with a cursor in the middle; the subject used the mouse to move the cursor to click on the location that he or she judged to be the target centroid location, i.e., the center of gravity of the locations of the disks d weighted by $|\alpha(d)|$. A feedback display (see Fig. 2 panel 4) was then presented which included (1) all the disks present in the stimulus presentation, (2) the subject's response location shown in blue, and (3) the location of the target centroid shown by a green bullseye. The subject could view the feedback for as long as desired. Pressing "Enter" on the keyboard initiated the next trial. Each subject performed 15 blocks of 100 trials.

In each of the Even and Odd tasks, the target location (indicated in the feedback display) had x and y coordinates

$$\begin{aligned}\text{target}_x &= \frac{1}{7}\sum_{\text{all disks } d}|\alpha(d)|\,C_x(d) \quad \text{and} \quad \text{target}_y \\ &= \frac{1}{7}\sum_{\text{all disks } d}|\alpha(d)|\,C_y(d)\end{aligned} \quad (4)$$

where $\alpha(d)$ is the α level of disk d, and $C_x(d)$ and $C_y(d)$ are the x and y coordinates of the center of disk d. The constant $\frac{1}{7}$ arises due the fact that there are two disks with each of the 7 levels of $\alpha = 0, \frac{1}{6}, \cdots, 1$; thus the total α-weight of all of the disks in any display is $2 \times \left(0 + \frac{1}{6} + \cdots + 1\right) = 7$.

3. Results

Throughout this section, we will present only those figures that are critical to understanding the main trends in the data; fine-grained details will be relegated to the supplementary material. We will refer to a figure in the supplementary materials as a "SupMatFig.".

All model fits reported below are derived using all 15 blocks of data provided by each subject. The results are not appreciably different if the first 5 blocks are discarded from the analysis for each subject.

3.1. The general modeling framework

We imagine that the visual system applies a battery of spatially

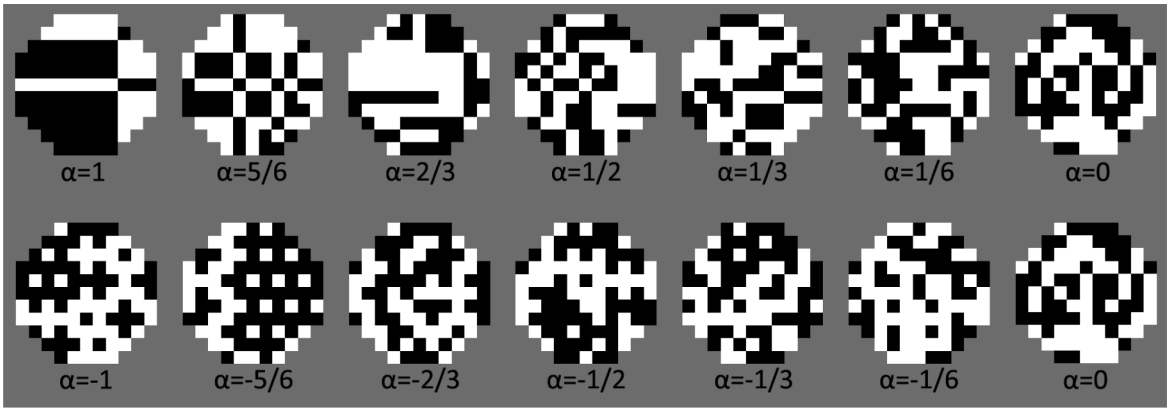

Fig. 3. *Stimulus texture disks*. The top (bottom) row gives a sample of the seven types of texture disks occurring in a given stimulus in the Even (Odd) condition. Each stimulus contains 14 texture disks, drawn either from the top row or the bottom row, with each type of texture in that row used twice.

parallel, roughly shift-invariant transformations to the visual input. We can think of the output from any one of these transformations as a "neural image" (Robson, 1980) whose pattern of activation reflects the distribution across visual space of some particular image statistic. We think of these neural images as the raw material available to the subject for use in tasks such as the Even and Odd tasks. We don't know what statistics are sensed by these up-front transformations, nor do we know how the subject is able to combine them. We assume, however, that when faced with either of the Even or Odd tasks, the subject uses top-down signals to synthesize a single task-specific, neural filter from these transformations. We think of this filter itself as a spatially local, roughly shift-invariant image transformation; that is, the filter assigns to each location (x, y) in space a real number that is a function of the visual input in some local neighborhood of (x, y). We call this neighborhood the "window" of the filter. Our aim is to discover the key properties of the filters used by our subjects in the Even and Odd tasks.

Basic questions addressed in our analysis are:

1. What is the window of the task-specific filter used by the subject in each of the Even and Odd tasks?
2. What is the relative sensitivity of each of these task-specific filters to the various different patterns that can appear within its window?

3.2. The α influence functions

The first step in our analysis is to infer the amount of influence that a disk d with a given level $\alpha(d)$ has on the centroid. This influence function, $W(\alpha)$, can be determined by linear regression of the response locations, against the locations of the individual disks on each trial. Specifically, for this analysis, we assume that on each trial, the x and y coordinates of the subject's response are given by

$$R_x = \frac{1}{2}\sum_{\text{disks } d} W(\alpha(d))C_x(d) + \beta_x + N_x \quad \text{and} \quad R_y = \frac{1}{2}\sum_{\text{disks } d} W(\alpha(d))C_y(d) + \beta_y + N_y, \tag{5}$$

where each sum ranges over all fourteen disks d in the stimulus, $C_x(d)$ and $C_y(d)$ are (as above) the x and y coordinates of the center of disk d, and

1. the function W: $\left\{0, \frac{1}{6}, \cdots, 1\right\} \to \mathbb{R}$ is normalized to sum to 1 (implying that the sum of $W(\alpha(d))$ over all disks d in any stimulus is equal to 2),
2. β_x and β_y are horizontal and vertical response biases fixed across trials, and
3. N_x and N_y are independent, normally distributed random variables with mean 0 and standard deviation σ.

Our interest is restricted to the function W, which gives the relative influence exerted on the subject's responses by disks d with different levels of α; accordingly, we will not report the other model parameters, β_x, β_y and σ.

The solid lines with error bars (95% Bayesian credible intervals) in Fig. 4 plot the influence exerted by disks d with different levels of $\alpha(d)$ on the responses of our subjects in the Even and Odd tasks. These influence functions are all normalized to sum to 1. The dashed lines plotted with triangles show the weighting function that was used to give feedback. Note that if the subject is accurately performing the centroid task, then the influence function $W(\alpha)$ will match the weighting function used for feedback, and be proportional to α.

The strong deviations of the influence functions from proportionality suggest that, despite the feedback, subjects were not able to respond in a manner that weighted the disks in proportion to α. In each task, every subject gave more weight to disks d with $|\alpha(d)| = 1$ than to disks with $\alpha(d) = 0$. In three or four cases, however, this difference is quite small. In these cases, subjects gave nearly equal weight to all disks d regardless of $\alpha(d)$ suggesting that they were able to attain very little traction in the task.

Note that in every case in which the α-influence function shows a large increase, the function tends to accelerate with $|\alpha(d)|$. That is, the α-influence function tends to be flat for values of $\alpha(d)$ near 0 and steepens in the Even task for values of $\alpha(d)$ near 1 and in the Odd task for values near -1. This observation implies immediately that the filters deployed by our subjects are not sensitive to α *per se*. If this were the case, then the α-influence functions achieved by our subjects would have to vary linearly with α.

3.3. Using block-filters for modeling

In all of the modeling we undertake below we assume that in a given task, on each trial the subject applies a fixed filter to the stimulus display and responds by mouse-clicking the centroid of the filter output (plus noise and a constant spatial bias). In particular, we will assume that the filter used by the subject is aligned to blocks of checks of a specific size and shape and responds differentially to different stimulus patterns falling within such blocks. By focusing on "block-filters" of this sort, we can gain insight into the nature of the window used in the filters deployed by our subjects.

Specifically, we assume the response location (R_x, R_y) on a given trial is given by

$$R_x = \frac{1}{T}\sum_{\text{blocks } b} F(\tau_b)x_b + \beta_x + N_x \quad \text{and} \quad R_y = \frac{1}{T}\sum_{\text{blocks } b} F(\tau_b)y_b + \beta_y + N_y, \tag{6}$$

where

1. the term "block" refers to a rectangular subpattern of the stimulus that contains only black or white checks (no gray checks). An $m \times n$ block comprises m rows by n columns of checks. To limit the number of free parameters required to model our results, we assume that the task-specific filters used by our subjects are invariant with respect to pattern rotation, reflection and contrast reversal. For example, the activation produced in a filter by a given $n \times m$ block b is identical to the activation produced by the $m \times n$ block produced by rotating b by 90°.
   Accordingly, in all of the models fit below, we choose specific integers m and n, and each of the sums in Eq. (6) is over all blocks b in the stimulus that are either $n \times m$ or $m \times n$. In the case in which $m \neq n$, we refer to the block-filter as an "$[m \times n \cup n \times m]$ block-filter."
2. τ_b is the "type" of block b: this is the equivalence class of block patterns that can be generated from b by some combination of rotation, reflection and/or contrast reversal. Thus, for example, there are four types of 2×2 blocks. These are shown in the upper panel of Fig. 5.
3. for any block-type τ, $F(\tau)$ is the activation produced in the filter by blocks of type τ.
4. x_b and y_b are the center x- and y-locations of block b. Thus, for example, each of the disks in the right panel of Fig. 5 falls in the center of a particular 2×2 block.
5. β_x and β_y are horizontal and vertical response biases independent of the stimulus. They are thus nuisance parameters of no real interest, but because they differ significantly from 0 for nearly all subjects, we include them to avoid distorting estimates of the parameters of interest. We shall not report their values.
6. N_x and N_y are normal random variables with mean 0 and some standard deviation σ, and
7. T gives the total activation produced in the filter by the current stimulus:

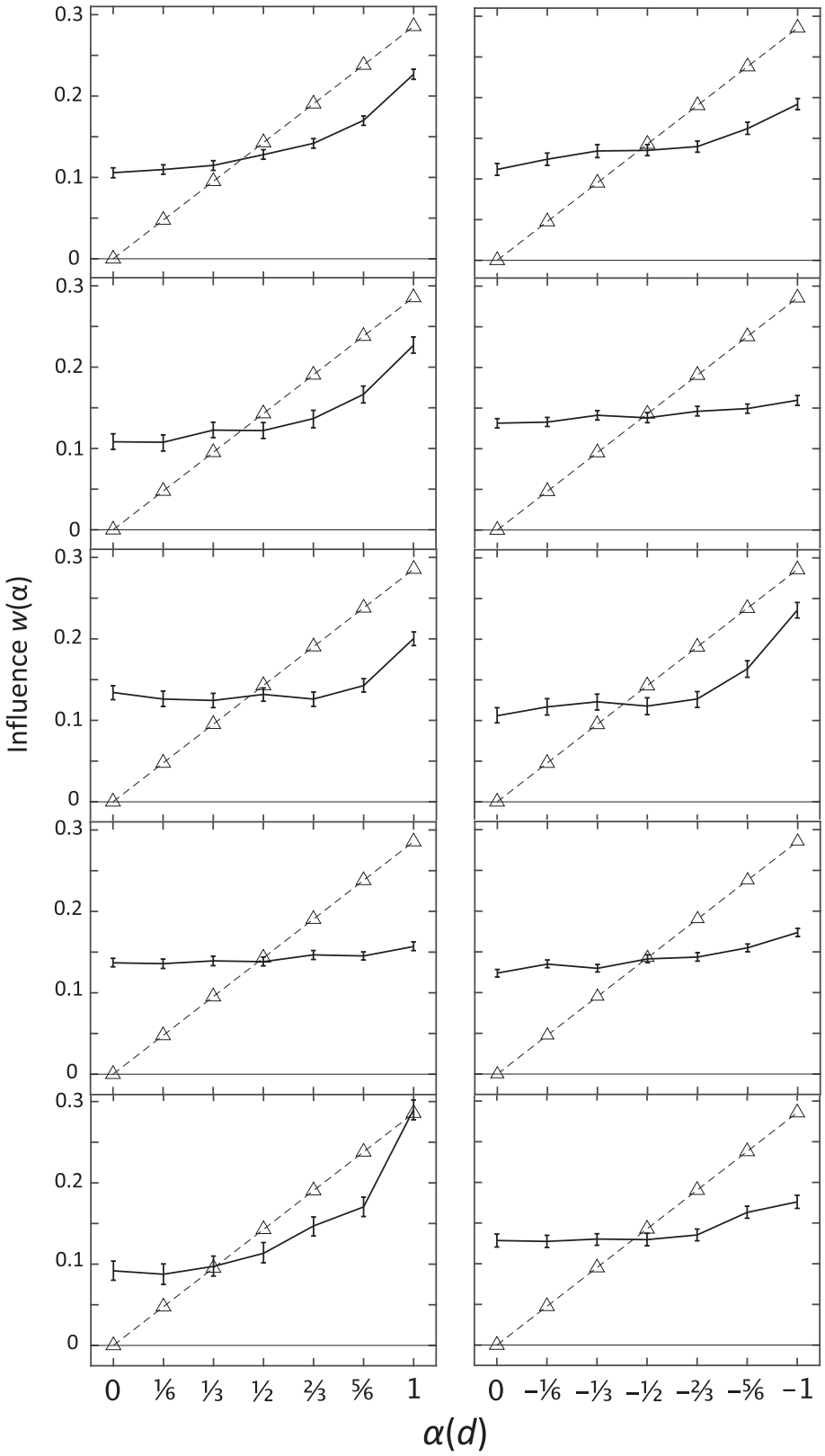

Fig. 4. *The α-influence functions.* Solid lines give the weight exerted by disks *d* with different values of $\alpha(d)$ on all subjects' centroid responses in the Even (left panels) and Odd (right panels) tasks. Error bars show 95% Bayesian credible intervals. The triangles show the target weights used to give feedback. Error bars give 95% Bayesian credible intervals.

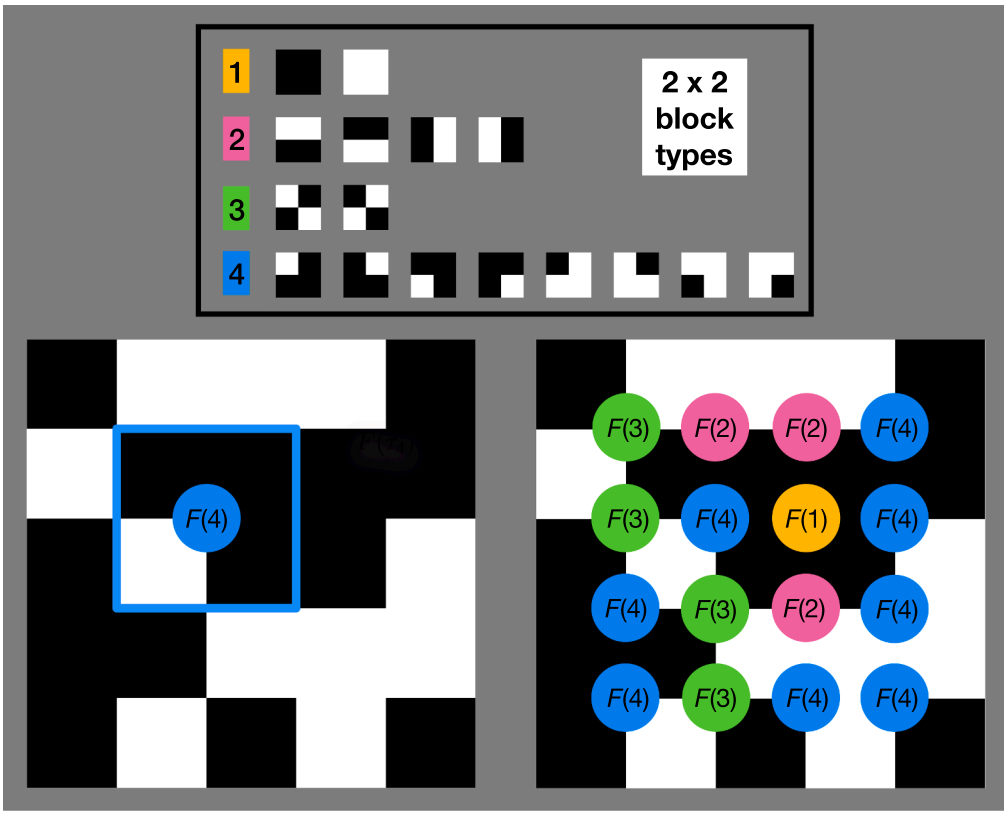

Fig. 5. 2 × 2 *block-filter.* The upper panel shows the four types of 2 × 2 blocks. Note that types 1, 2 and 3 have even parity, and type 4 has odd parity. The left lower panel shows the window (blue line) used to assign the value $F(4)$ to the central location of the 2 × 2 block captured by the window. The right panel shows all of the values assigned within the patch by the block-filter. Each location receiving a value is the center of a 2 × 2 block of checks, and the value assigned depends on the type of the block.

$$T = \sum_{\text{blocks } b} F(\tau_b). \tag{7}$$

The reader will note that the model of Eq. (6) is underconstrained in the sense that the function F is defined only up to an arbitrary multiplicative constant. We deal with this by imposing the constraint that

$$\sum_{\text{all block-types}\tau} F(\tau)^2 = 1. \tag{8}$$

3.4. *Is it possible that our subjects' task-specific filters are* 2 × 2 *block-filters?*

The first possibility that we consider is that the subject processes 2 × 2 regions of the image, but uses a filter whose output is not the veridical value of alpha. That is, we ask whether it is possible to account for the subject's responses in terms of a 2 × 2 block-filter. There are four types of 2 × 2 blocks. These four types are shown in the top panel in Fig. 5. Note that types 1, 2 and 3 all have even parity and type 4 has odd parity.

The left lower panel in Fig. 5 shows how a single 2 × 2 block produces an output value. One can think of the output produced by a particular 2 × 2 block b of type τ_b as a region of activation of total weight $F(\tau_b)$ centered at the location in the middle of b. In the lower left panel of Fig. 5, the blue line shows the window encompassing a single 2 × 2 block of checks. Because this particular 2 × 2 block is of type 4, the filter assigns the value $F(4)$ to the central location of the 2 × 2 block captured by the window. The right panel shows all of the values assigned within the patch by the filter. Each location receiving a value is the center of a 2 × 2 block of checks, and the value assigned depends on the type of the block. The model's prediction of the subject's response is given by Eq. (6) where the sums are over all 2 × 2 blocks b.

3.4.1. *The estimated* 2 × 2 *block-filters*

SupMatFig. 1 shows the 2 × 2 block-filters $F(\tau)$ estimated for all five subjects in each of the Even and Odd tasks under the assumption that responses are produced by Eq. (6).

For all cases except those in which the observed α-influence function is flat, our results decisively reject the possibility that our subjects are using 2 × 2 block-filters. We can see this by using the 2 × 2 block-filters shown in SuppMatFig. 1 to predict the weights exerted on the responses of our subjects by disks d with different levels of $\alpha(d)$. These predicted weights are plotted by the gray circles in SupMatFig. 2.

Here we focus on the single example of the α-influence function observed for subject S5 in the Even task. The steep acceleration of this particular α-influence function make this the most stringent modeling challenge we face for any subject in either the Even or Odd task.

The α-influence function predicted under the assumption that S5 is using a 2 × 2 block-filter in the Even task (with the block-filter given by the bottom left panel of SupMatFig. 1) is shown by the gray disks in the upper left panel of Fig. 6. (The white-filled squares replot the estimated α-influence function from the bottom left panel of Fig. 4.) As is clear, the observed α-influence function deviates strongly from the function predicted from the 2 × 2 block-filter model.

In fact, the mode of failure of the 2 × 2 model seen in the upper left panel of Fig. 6 is generic: any 2 × 2 block filter model will produce a linear influence function. This is a consequence of the fact that the probability of occurrence of any given 2 × 2 block pattern varies linearly with α (in some cases increasing, in some cases decreasing), so any linear function of these probabilities (given by the block-filter activations) will necessarily be linear as well. The model of Eq. (6) thus

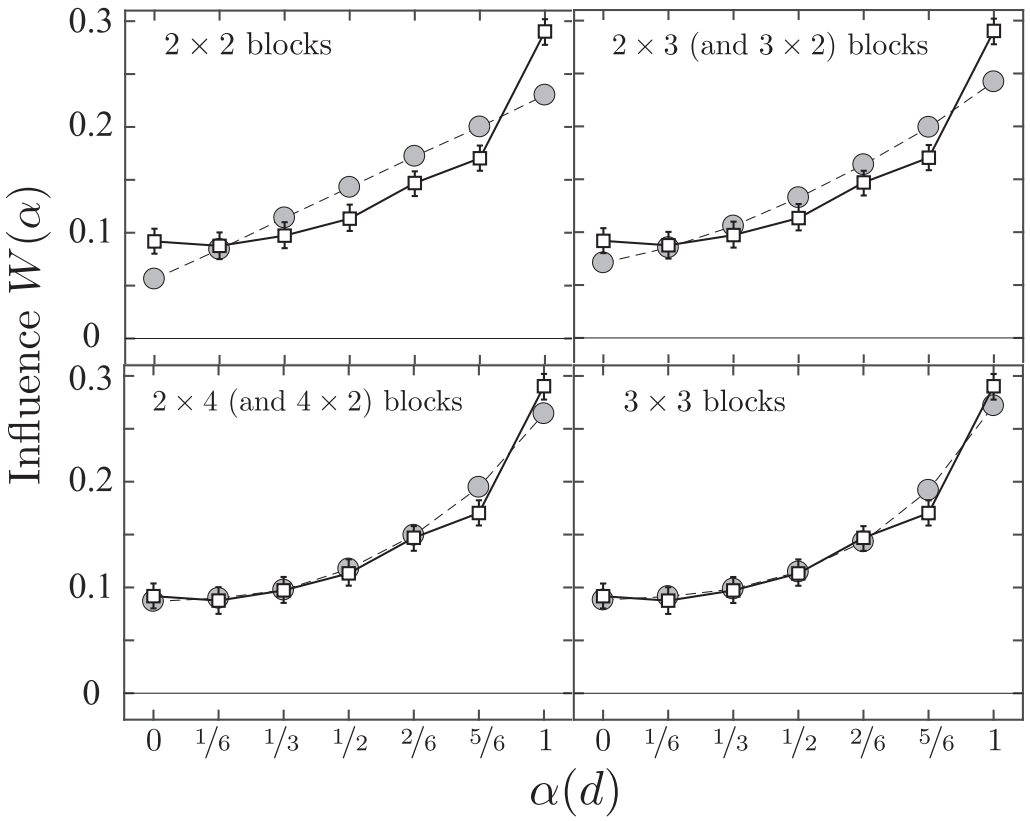

Fig. 6. *Predicted α-influence functions derived from fits of different block-filter models for subject S5 in the Even task.* White-filled squares with error bars (95% Bayesian credible intervals) replot the α-influence function from the bottom left panel of Fig. 4. Gray-filled circles with dashed lines plot predicted weights $W(\alpha)$ exerted by disks with different levels of α under the assumption that in the Even task S5 is using a 2 × 2 block-filter (upper left panel), a [2 × 3 ∪ 3 × 2] block-filter (upper right panel), a [2 × 4 ∪ 4 × 2] block-filter (lower left panel), and a 3 × 3 block-filter (lower right panel).

implies that the weights assigned different disks d must vary linearly with $\alpha(d)$ regardless of the activations $F(\tau)$ produced by different 2 × 2 pattern-types τ.

Subject S4 showed very little differential sensitivity to disks varying in α in either the Even or the Odd task. This is also true of subject S2 in the Odd task. Unsurprisingly (as revealed by SupMatFig. 2), in these cases the linear α-influence functions predicted assuming 2 × 2 block-filters match the observed α-influence functions very well. However, in all other cases, the predicted α-influence functions deviate strongly from the accelerated form of the observed α-influence functions.

3.5. Block-filters larger than 2 × 2 can achieve accelerated α-influence functions.

The heightened sensitivity of the filters achieved by our subjects to disks d with $|\alpha(d)| = 1$ suggests that these filters are selectively activated by features that emerge with high probability predominantly in texture patches when $|\alpha|$ nears 1. For example, large homogeneous white (or black) patches are relatively common in pure even texture but rare in texture patches d with $\alpha(d) \leqslant \frac{5}{6}$.

Beyond the 2 × 2 size, the probability of occurrence of blocks typically depends in a nonlinear fashion on α. Fig. 7 illustrates this effect for 3 × 3, 4 × 3 and 4 × 4 blocks comprising either all black or all white checks. The top panel shows the average number of occurrences in texture disks from the Even task of solid, 3 × 3 blocks (squares), 4 × 3 (or 3 × 4) blocks (triangles) and 4 × 4 blocks (circles). In the bottom panel, each of the curves from the top panel is rescaled to have sum equal to 1; this reveals the increasing acceleration of the curve for 4 × 4 blocks vs. 4 × 3 vs. 3 × 3 blocks.

However, sensitivity to solid blocks is unlikely to account for our findings. As shown in Fig. 7, very few solid, 4 × 4 blocks occur in any of the texture disks used in the Even task. In particular, even disks d with the maximum value of $\alpha(d) = 1$ contain on average only 0.62 solid, 4 × 4 blocks. In fact, the rate of occurrence of solid, 4 × 4 blocks is sufficiently low that on 5% of all trials, all fourteen disks in the stimulus will be entirely devoid of solid, 4 × 4 blocks. Thus, performance in the Even task would be very poor if one were to rely exclusively on solid, 4 × 4 blocks to guide one's responses. By contrast, a filter selective in varying degrees for all solid $n \times m$ blocks ($n > 2$, $m > 2$) could well provide enough information to enable effective performance in the Even task.

The second thing to note about Fig. 7 is that all of the lower-panel curves for solid 3 × 3, 4 × 3 and 4 × 4 blocks assign values close to 0 to disks d with $\alpha(d) = 0$. This implies that if the filters achieved by our subjects were purely selective for large, solid blocks of checks (i.e., if they consisted of linear combinations of the curves plotted in the top panel of Fig. 7), then they would necessarily assign values close to 0 to disks d with $\alpha(d) = 0$. By contrast, as shown by Fig. 4, the influence functions achieved by our subjects all give weight substantially greater than 0 to disks d with $\alpha(d) = 0$. We conclude that the filters achieved by our subjects are not purely selective for large, solid blocks of checks; these filters must also be selective for features that occur with moderately high probability in disks d with $\alpha(d) = 0$.

3.6. What size block-filter is required to explain the current data?

In this section, we will consider in succession the possibility that our data can be explained by [2 × 3 ∪ 3 × 2], [2 × 4 ∪ 4 × 2], and 3 × 3 block-filters. In each case, we will start by fitting the model of Eq. (6) to estimate the block-filter $F(\tau)$ used by each subject. We will then use the estimated block-filter to predict the subject's α-influence function.

When we plot the block-filter $F(\tau)$ estimated for a given subject for

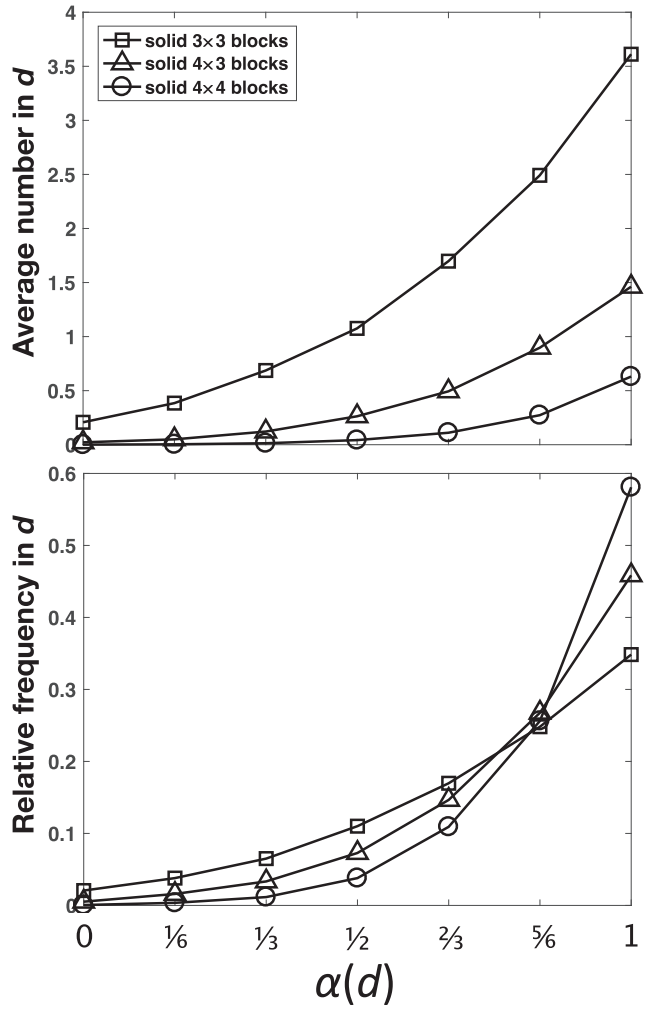

Fig. 7. *Rate of occurrence of larger, homogeneous blocks in Even-task texture disks.* For disks d with average levels of $\alpha = 0, \frac{1}{6}, \cdots, 1$, the top panel gives the average number of occurrences of homogeneous blocks (either all black or all white) comprising 3×3 checks (circles), 4×3 (or 3×4) checks (triangles) or 4×3 checks (squares).

a given block size, it will be useful to indicate the evenness of all of the different block-types τ. For a given $n \times m$ block b, we can define $\alpha(b)$ by using b in place of d in Eq. (1), where $p(b)$ gives the proportion of 2×2 subblocks of b that have even parity. As is easy to check, all blocks b of a given block-type τ the same value of $\alpha(b)$; we will write $\alpha(\tau)$ for this value.

3.6.1. Can the data be explained in terms of $[2 \times 3 \cup 3 \times 2]$ *block-filters?*

As noted above, the results plotted in Fig. 4 are inconsistent with the idea that our subjects are using 2×2 block-filters. Can these results be modeled in terms of $[2 \times 3 \cup 3 \times 2]$ block-filters? In testing this possibility, we assume specifically that responses are produced in accordance with Eq. (6) with each sum ranging over all blocks b in the stimulus that are either 2×3 or 3×2.

There are 14 different types of 2×3 (or 3×2) blocks. These are shown in SupMatFig. 3. The estimated $[2 \times 3 \cup 3 \times 2]$ block-filters $F(\tau)$ are shown in SupMatFig. 4 for all subjects in both the Even and Odd tasks. To show how the different block-types vary in evenness, the light gray open circles in SupMatFig. 4 plot $0.4\alpha(\tau)$.

As shown by the upper right panel of Fig. 6, the α-influence function observed for subject S5 in the Even task deviates strongly from the α-influence function predicted assuming the subject is using a $[2 \times 3 \cup 3 \times 2]$ block-filter. More generally, as shown by SupMatFig. 5, in all cases in which the curve from Fig. 4 is accelerated, the curve predicted using the estimated $[2 \times 3 \cup 3 \times 2]$ block-filter, while accelerating, fails to match the data in steepness.

3.6.2. Can the data be explained in terms of $[2 \times 4 \cup 4 \times 2]$ *block-filters?*

Here we assume that responses are derived from Eq. (6) with each sum ranging over all blocks b in the stimulus that are either 2×4 or 4×2. The 44 different types τ of 2×4 (or 4×2) blocks are shown in SupMatFig. 6 sorted into an order in which $\alpha(\tau)$ is non-increasing.

The estimated block-filter activations are shown in SupMatFig. 7 for all subjects in both the Even and Odd tasks. The light gray open circles in SupMatFig. 7 plot $0.4\alpha(\tau)$ for all block-types τ. The darkness of a marker reflects the proportion of all 2×4 and 4×2 blocks occurring in all stimuli that are of the corresponding type. Thus, block-types with low indices (indicating greater evenness) in SupMatFig. 6 are more common in the Even task than block-types with high indices, and vice versa for the Odd task. Markers whose error bars (99% Bayesian credible intervals) that do not contain 0 have thickened bars.

SupMatFig. 8 replots the curves from Fig. 4 (white markers) along with the weights that disks with different levels of α would be predicted to exert if these weights were determined by only the 2×4 and 4×2 blocks in the disks, where each block is assumed to contribute the weight shown in SupMatFig. 7. In most cases, the predicted curve matches the curve from Fig. 4 fairly well. The one deviation is for subject S5 in the Even task (also plotted in the lower left panel of Fig. 6); in this case, the predicted curve overshoots the observed influence exerted by disks with $\alpha = \frac{5}{6}$ and undershoots the influence exerted by disks with $\alpha = 1$.

3.6.3. Can the data be explained in terms of 3×3 *block filters?*

SupMatFig. 9 displays all 51 types of 3×3 block patterns sorted into an order in which $\alpha(\tau)$ is non-increasing.

The relative weights $F(\tau)$ exerted by the 51 block types τ are shown in SupMatFig. 10 for all subjects in both the Even and Odd tasks. Points on the horizontal axis correspond to the 51 types of 3×3 blocks shown in SupMatFig. 9. Plotted values reflect the weights exerted by different block-types on the responses of all five subjects (rows 1–5) in the Even and Odd tasks (left and right columns). The light gray open circles show $0.4\alpha(\tau)$ for all 51 block types τ.

The darkness of a plotted point reflects the relative number of times patterns of that type appeared in the stimuli. As one might expect, the error bars (99% Bayesian credible intervals) tend to be smaller for more common 3×3 block-types. Bars are thickened for block-types τ for which the error bars on $F(\tau)$ do not include 0.

The α-influence functions plotted in Fig. 4 are consistent with the idea that our subjects are using 3×3 block-filters. This is shown by the dashed lines with gray disks in SupMatFig. 11. The dashed-line curve in a given panel gives the predicted influence exerted by disks varying in α-level on our subject's responses if the subject were using the 3×3 block-filter $F(\tau)$ in the corresponding panel of SupMatFig. 10. In particular, as shown in the lower right panel of Fig. 6, the observed α-influence function for subject S5 in the Even task is well-described by the α-influence function predicted assuming a 3×3 block-filter.

3.7. What features of the Even and Odd textures activate subject filters most strongly?

The observed α-influence functions (Fig. 4) reject the possibility that our subjects are using either 2×2 or $[2 \times 3 \cup 3 \times 2]$ block-filters. However, the α-influence functions predicted assuming subjects are using either $[2 \times 4 \cup 4 \times 2]$ or 3×3 block-filters match the observed α-influence functions fairly well.

This suggests that we may be able to gain insight into the features that operate most strongly to activate the filters used by our subjects in both tasks by focusing on the particular $[2 \times 4 \cup 4 \times 2]$ and 3×3 block-types that predominate in controlling performance in the Even and Odd

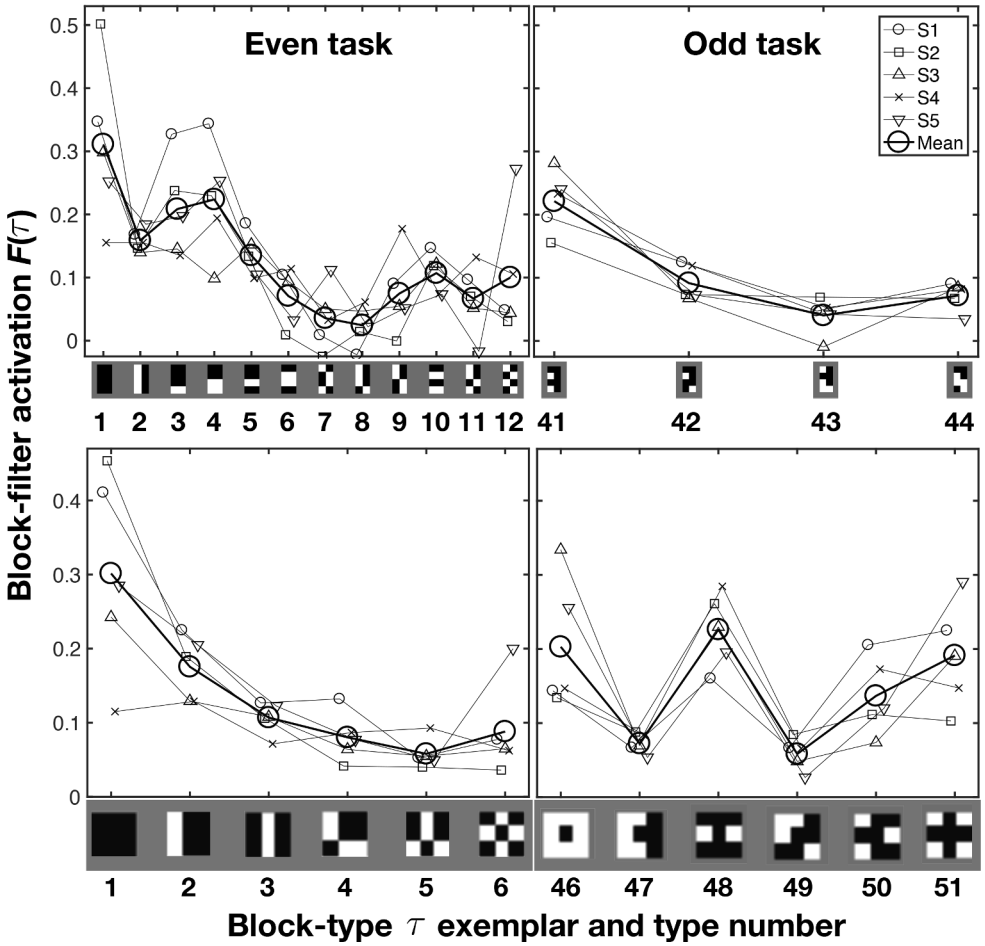

Fig. 8. *Sensitivity of* [2 × 4 ∪ 4 × 2] *and* 3 × 3 *block-filters to the most influential block types.* Each curve in the upper left (right) panel gives the [2 × 4 ∪ 4 × 2] block-filter observed for a single subject in the Even (Odd) task across the 12 (4) block-types whose three 2 × 2 subblocks all have even (odd) parity. Each curve in the lower left (right) panel gives the [2 × 4 ∪ 4 × 2] block-filter observed for a single subject in the Even (Odd) task across the 6 block-types whose four 2 × 2 subblocks all have even (odd) parity.

tasks because they occur most frequently in the stimuli.

3.7.1. Features controlling performance in the Even task

In the stimuli used in the Even task, the most commonly occurring [2 × 4 ∪ 4 × 2] and 3 × 3 block types are the ones whose 2 × 2 subblocks all have even parity. There are 12 such [2 × 4 ∪ 4 × 2] block-types and 6 such 3 × 3 block-types. The [2 × 4 ∪ 4 × 2] and 3 × 3 block-filters estimated for all five subjects in the Even task across these block types are plotted in the upper and lower left panels of Fig. 8 respectively. In each panel, the large circles plot the mean block-filter.

The 3 × 3 and [2 × 4 ∪ 4 × 2] block-filters for the Even task conform to a simple rule. The solid blocks (i.e., blocks of type 1) produce the highest activation; blocks comprising two subregions (3 × 3 blocks of type 2, and [2 × 4 ∪ 4 × 2] blocks of types 2, 3 and 4) exert the next highest activation; and all block-types comprising more than 2 subregions tend to produce lower and roughly equal levels of activation.

3.7.2. Features controlling performance in the Odd task

In the stimuli used in the Odd task, the most commonly occurring [2 × 4 ∪ 4 × 2] and 3 × 3 block types are the ones whose 2 × 2 subblocks all have odd parity. There are 4 such [2 × 4 ∪ 4 × 2] block-types and 6 such 3 × 3 block-types. The [2 × 4 ∪ 4 × 2] and 3 × 3 block-filters estimated for all five subjects in the Odd task across these block types are plotted in the upper and lower right panels of Fig. 8 respectively.

The [2 × 4 ∪ 4 × 2] block-filters for the Odd task are easy to characterize. [2 × 4 ∪ 4 × 2] blocks of type 41 produce the highest activation; [2 × 4 ∪ 4 × 2] blocks of types 42, 43, and 44 produce lower and roughly equal activation.

The 3 × 3 block-filters for the Odd task are more complicated. For all subjects, 3 × 3 block-type 48 produces activation that is clearly greater than either of block-types 47 and 49. There is more variability across subjects in the activation produced by block-types 46, 50 and 51.

The 3 × 3 block-types that produce the highest activation have greater symmetry than the other three. The three 3 × 3 block-types that produce the highest average activation are 46, 48 and 51. Block-types 46 and 51 are symmetric with respect to reflection across the vertical, horizontal and both diagonal axes; block-type 48 is symmetric with respect to vertical and horizontal reflection. By contrast, each of the 3 × 3 block-types 47, 49 and 50 is symmetric with respect to reflection around a single axis. In the cases of block-types 47 and 50, this is the horizontal axis; in the case of 49, this is the top-left-to-bottom-right diagonal axis.

4. Discussion

The motivation for this work is to understand the computations underlying the extraction of high-order statistics, since, as is well-known, these statistics are critical for conveying visual features, such as corners and edges (Morrone et al., 1982; Oppenheim & Lim, 1981). We focus on the four-point correlation among points in a 2 × 2 neighborhood. This correlation is visually salient when isolated in synthetic visual textures (Julesz et al., 1978) and has come to be known as even-

ness, or α (Victor & Conte, 1991, 2012; Victor et al., 2005). For black-and-white textures, α is defined as the average value of the product of the signed contrasts in four checks in a 2×2 neighborhood. Textures defined by α are readily distinguished from random textures (Victor, Thengone, Risvi, & Conte, 2015), and it is known that neurons in V1 and especially V2 can respond to this distinction as well (Yu et al., 2015). But four-point multiplication is not a computation generally thought to be carried out by visual neurons. This discordance suggests that visual neurons do not compute α *per se*, but instead are sensitive to aspects of visual structure that α implies. Such alternatives had been previously proposed (Victor & Conte, 1989, 1991) but not tested experimentally. Here we carry out such a test. Results show that α is in fact not directly computed, and begin to characterize the computations that indirectly determine it.

Our strategy was motivated by the statistical fluctuations of image statistics that inevitably occur in small patches of texture. In an adaptation of the "centroid paradigm" (Drew et al., 2010; Sun et al., 2016), we asked subjects to determine the centroid of α (or $-\alpha$) in an array of 14 small texture disks and (to exaggerate these fluctuations) constructed the disks so that the precise values of $|\alpha|$ within each disk varied from 0 to 1. As in Drew et al., 2010 and Sun et al., 2016, we modeled the subjects' responses in terms of how the values of α in each disk influenced their decision. This showed that even after extensive feedback based on the veridical value of $|\alpha|$, subjects chose the centroid based on an influence function that was an accelerating function of $|\alpha|$, rather than proportional to it. We note that in a centroid paradigm based on item gray-scale or item contrast, subjects perform in a veridical fashion (Drew et al., 2010; Sun et al., 2016); this indicates that their failure to do so here is due to the visual computations applied to the stimulus disks, not the task itself. This observation ruled out the possibility that subjects directly computed α. It also ruled out the possibility that subjects were carrying out some other nonlinear computation within 2×2 neighborhoods and using the summed value of this computation rather than α-as this would also have led to a linear α-influence function. Rather, an accelerating influence function could only be explained by computations in which the contrasts in neighborhoods larger than 2×2 interacted in a nonlinear way. Nonlinearities that took into account 2×3 regions could provide for a quadratic acceleration (since they contain two 2×2 regions), but the observed acceleration required still larger regions- a minimum of 2×4 or 3×3. These findings are consistent with the findings of Victor and Conte, 1989, who used a VEP analysis to infer that correlations within a $2 \times n$ region ($n \approx 4$) were combined in a nonlinear fashion. In this connection, it should be noted (Tkačik, Prentice, Victor, & Balasubramanian, 2010) that human observers are sensitive only to those higher-order correlations that vary informatively in natural scenes, and these correlations correspond to the particular 4-check subsets of 3×3 blocks that have been identified as being highly salient by Victor and Conte, 1991. Taylor et al., 2008 looked at psychophysical sensitivity to a wide variety of local statistical structures, and concluded that human performance could be accounted for mechanisms that processed "samplers" containing up to 10 pixels, arranged in parallel strips. Interestingly, their strategy showed that this is an upper limit of what is needed to account for performance; the present work shows that for the high-order correlation defined by α, "samplers" of 2×4 or 3×3 pixels is a lower limit. In an interesting reanalysis of the data of Taylor et al., 2008; Nagai et al., 2009 were able to account for the observed human sensitivity to the texture differences in terms of recursive, nonlinear processing accomplished by oscillator networks that included as few as four oscillators.

The work of Barbosa et al., 2013 sheds an interesting light on the relationship between features that can be computed within a 2×2 block and features that are computed from larger blocks. They present an account of texture discrimination in which the third Minkowski functional (Michielsen & De Raedt, 2001) plays a key role. This functional describes the frequency of "holes" in a texture (in which the black checks are interpreted as material, and the white checks as empty spaces). At first glance, the topological nature of this feature suggests that, in order to calculate this quantity, the visual system might need to examine large blocks of texture. However, as they have shown, this functional can actually be calculated from 2×2 blocks. But, as we have shown here, an account of performance in the centroid task requires that larger blocks are taken into account.

The present analysis gives a second level of insight via the block-filters, as they indicate which kinds of 2×4 or 3×3 blocks are the primary contributors to the local computation. As seen in Fig. 8 (and also in SupMatFigs. 7 and 10), these configurations tend to have homogeneous regions and high spatial symmetry- though we stress that neither homogeneity alone nor symmetry alone accounts for the influential block types.

We should note, however, one shortcoming of the centroid method. It has often been observed that subjects are more sensitive to differences in α than to differences in $-\alpha$. In particular, in tasks requiring subjects to judge the location of a target in a background of binary texture with $\alpha = 0$, if the target is defined by α, then the threshold is roughly 25% lower than it is if the target is defined by $-\alpha$ (Victor et al., 2015). Although the centroid method used here reveals the *relative* influence of texture features in controlling our subjects' responses in the Even and Odd tasks, this approach does not bring out overall sensitivity differences to variations in α vs. $-\alpha$ shown clearly in other studies.

Finally, we mention that the texture centroid paradigm introduced here provides a general strategy for analyzing how visual image statistics are processed. By determining the weights that subjects actually use to represent the characteristics of texture patches, it can be adapted to explore the computations used to process other kinds of features that drive human discrimination, such as other varieties of isodipole texture, colored textures, or natural textures (Portilla & Simoncelli, 2000). As in the present application, we anticipate that the centroid paradigm will provide two levels of information: first, the influence function, which captures the overall relationship between the amount of a feature and its impact on the centroid; and second, the "block filters," which characterize the local analysis that computes this feature.

Acknowledgments

This research was supported by NIH Grant EY07977 to Jonathan D. Victor.

Appendix A

A.1. Concerning the Bayesian model-fitting procedure.

Here we describe the details of the Bayesian method used to estimate the joint posterior density characterizing the parameters of each of the two models used in the paper. It will be convenient to

1. The model of Eq. (5) has parameters σ and $V = \left(W(0), W\left(\frac{1}{6}\right), \cdots, W(1), \beta_x, \beta_y\right)$.
2. The model of Eq. (6) has parameters σ and $V = (F(\tau_1), F(\tau_2), \cdots, F(\tau_N), \beta_x, \beta_y)$, where N is the number of block-types, and τ_k is the k^{th} block-type.

A.1.1. The likelihood functions for the two models

As indicated above, we write V for the vector of parameters other than σ for the model in question. We also write $\phi(x, \sigma)$ for the normal pdf with mean 0 and standard deviation σ: i.e.,

$$\phi(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-x^2}{2\sigma^2}} \tag{9}$$

Then the likelihood function (in practice we compute the log of this function) is

$$\Lambda(V, \sigma) = \prod_{\text{all trials}t} \phi(R_{t,x} - \widehat{R}_{t,x}(V), \sigma)\phi(R_{t,y} - \widehat{R}_{t,y}(V), \sigma) \tag{10}$$

where, for $u = x, y$, $R_{t,u}$ gives the u-coordinate of the subject's response on trial t, and $\widehat{R}_{t,u}(V)$ is the expectation of this value given the parameter values in V. I.e., in the case of the model of Eq. (5),

$$\widehat{R}_{t,u}(V) = \frac{1}{2} \sum_{\substack{\text{all disks } d \\ \text{on trial } t}} W(\alpha(d))C_u(d) + \beta_u \tag{11}$$

where (as in Eq. (5)) $\alpha(d)$ is the α-level of disk d and $C_u(d)$ is the u-coordinate of the center of disk d; and in the case of the model of Eq. (6),

$$\widehat{R}_{t,u}(V) = \frac{1}{T_t} \sum_{\substack{\text{all blocks } b \\ \text{on trial } t}} F(\tau_b)u_b + \beta_u \quad \text{with} \quad T_t = \sum_{\substack{\text{all blocks } b \\ \text{on trial } t}} F(\tau_b) \tag{12}$$

where (as in Eq. (6)) τ_b is the type of block b and u_b is the u-coordinate of the center of block b.

A.1.2. Markov chain Monte Carlo simulation

The estimation method uses Markov chain Monte Carlo (McMC) simulation. For simplicity, we use uniform prior distributions on all parameters. In any McMC process, one starts with some arbitrary guess at the parameter vector Q (which will ultimately be thrown away) and sets ${}_1S = Q$. (Note: (1) We use pre-subscripts to indicate parameter vector sample number in the McMC process; (2) In the current applications of this method, Q comprises guesses at σ as well as the parameters in V for whichever model is being fit.) Then one iterates the following steps some large number N_{iter} of times:

1. Pick a candidate parameter vector C in the neighborhood of the last sample ${}_{n-1}S$ added to the list. Then
2. for

$$R = \frac{\Lambda(C)f_{prior}(C)}{\Lambda({}_{n-1}S)f_{prior}({}_{n-1}S)} = \frac{\Lambda(C)}{\Lambda({}_{n-1}S)} \tag{13}$$

(where the righthand equality follows from the fact that we use uniform priors on all parameters).

- if $R \geqslant 1$, set ${}_nS = C$;
- otherwise set

$${}_nS = \begin{cases} C & \text{with probability } R \\ {}_{n-1}S & \text{with probability } 1 - R \end{cases}. \tag{14}$$

The classical result (Hastings, 1970) is that (provided the procedure for selecting candidates C satisfies certain conditions) in the limit as $N_{iter} \to \infty$ this algorithm yields a sample from the posterior density. In practice, one typically throws away the first several thousand samples from the list which are usually not representative of the samples accumulated after the McMC process has stabilized.

A.1.3. Priors

The bounds of the uniform priors matter very little provided they are wide enough to include the posterior density. In the current simulations, the prior density of each parameter other than σ was taken to be uniform between -10 and 10; the prior on σ was uniform between 0 and 100.

A.1.4. Adaptive candidate selection

The sampling window used to select the candidate parameter vector C on each iteration of the McMC process dramatically influences the efficiency with which one can estimate the posterior joint density of the parameters. We use an adaptive procedure to adjust the sampling window after every 1000 iterations of the process.

On any given iteration, we generate the candidate C from the last sample (${}_{n-1}S$) by first setting

$$\widetilde{C} = {}_{n-1}S + X \tag{15}$$

where X is a vector of jointly independent Gaussian random variables whose k^{th} entry (corresponding to the k^{th} model parameter) has standard deviation γ_k. (It is the γ_k's that are adjusted after each block of 1000 iterations.) In the case of the model of Eq. (5), $\widetilde{C}$ includes values $\widetilde{W}(0), \widetilde{W}\left(\frac{1}{6}\right), \cdots, \widetilde{W}(1)$; we generate the candidate vector C from $\widetilde{C}$ by dividing the values $\widetilde{W}(a)$ values by their sum (to impose the constraint that the values $W(a)$ must sum to 1). In the case of the model of Eq. (6) (writing N for the number of different block-types, and τ_k the k^{th} block-type), $\widetilde{C}$ includes values $\widetilde{F}(\tau_1), \widetilde{F}(\tau_2), \cdots, \widetilde{F}(\tau_N)$; we generate the candidate vector C from $\widetilde{C}$ by dividing the values $\widetilde{F}(\tau_k)$ by the square root of the sum of their squared values (to impose the constraint of Eq. (8)).

After a given block of 1000 iterations, we compute the standard deviation ρ_k of the kth model parameter across the previous 1000 iterations. If all

of the ρ_k's are 0 (suggesting that no candidate C was ever accepted over the course of the previous 1000 iterations and hence that we have been drawing candidates sufficiently far from ${}_{n-1}S$ that their likelihoods have all been much smaller than that of ${}_{n-1}S$), then we decrease all σ_k's by 10%. Otherwise, we adjust the γ_k's by setting $\gamma_k = A\rho_k$ where A is a scalar that was initialized to 0.1 and is itself adjusted after each 1000 trials according to the following rule: If the median value (across the previous 1000 trials) of the likelihood ratio R (Eq. (13)) is less than 0.5, A is decreased by 10%; otherwise A is increased by 10%. By adjusting A in this way, we insure that the median likelihood ratio R of Eq. (13) will be near 0.5 which guarantees that the process moves efficiently to scribble in the joint posterior density.

A.1.5. Starting values, burn-in, and number of iterations

In fitting the model of Eq. (5), $W(a)$ was initialized to $\frac{1}{7}$ for $a = 0, 1, \cdots, 6$; β_x and β_y were initialized to 0; and σ was initialized to 15. In fitting the model of Eq. (6), for all block-types τ, $F(\tau)$ was initialized to $\frac{1}{\sqrt{N}}$ for N equal to the number of different block-types, β_x and β_y were initialized to 0, and σ was initialized to 15. 120,000 iterations of the McMC process were performed, the first 20,000 of these were used to allow the McMC process to "burn in," and the last 100,000 were taken as a representative sample of the posterior joint density characterizing the parameters. In each case, the last 100,000 samples were plotted and inspected by eye to insure stability of parameter estimates. Stability of the obtained estimates was confirmed to hold across several other starting locations.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.visres.2019.03.006.

References

Barbosa, M. S., Bubna-Litic, A., & Maddess, T. (2013). Locally countable properties and the perceptual salience of textures. *Journal of the Optical Society of America A, 30*, 1687–1697.

Drew, S., Chubb, C., & Sperling, G. (2010). Precise attention filters for weber contrast derived from centroid estimations. *Journal of Vision, 10*(20), 1–16.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika, 57*, 97–109.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, 290*, 91–97.

Julesz, B., Gilbert, E. N., & Victor, J. D. (1978). Visual discrimination of textures with identical third order statistics. *Biological Cybernetics, 31*, 137–149.

Maddess, T., & Nagai, Y. (2001). Discrimination of isotrigon textures. *Vision Research, 41*, 3837–3860.

Michielsen, K., & De Raedt, H. (2001). Integral-geometry morphological image analysis. *Physics Reports-Review Section of Physics Letters, 347*(6), 462–538.

Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. i. Neurophysiological evidence. *Proceedings of the Royal Society of London B: Biological Sciences, 216*(1204), 335–354.

Nagai, Y., Taylor, R. R. L., Loh, Y-W., & Maddess, T. (2009). Discrimination of complex form by simple oscillator networks. *Network: Computation in Neural Systems, 20*(4), 233–252.

Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE, 69*(5), 529–541.

Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision, 40*, 49–71.

Purpura, K. P., Victor, J. D., & Katz, E. (1994). Striate cortex extracts higher-order spatial correlations from visual textures. *Proceedings of the National Academy of Sciences, 91*(18), 8482–8486.

Robson, J. G. (1980). Neural images: The physiological basis of spatial vision. In C. S. Harris (Ed.). *Visual coding and adaptability* (pp. 177–214). Hillsdale, NJ: Erlbaum.

Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2016). The centroid paradigm: Quantifying feature-based attention in terms of attention filters. *Attention, Perception and Psychophysics, 78*, 474–515. https://doi.org/10.3758/s13414-015-0978-2.

Taylor, R. R. L., Maddess, T., & Nagai, Y. (2008). Spatial biases and computational constraints on the encoding of complex local image structure. *Journal of Vision, 8*(7), 1–13 19.

Tkačik, G., Prentice, J. S., Victor, J. D., & Balasubramanian, V. (2010). Local statistics in natural scenes predict the saliency of synthetic textures. *Proceedings of the National Academy of Sciences, 107*(42), 18149–18154.

Victor, J. D. (1986). Isolation of components due to intracortical processing in the visual evoked potential. *Proceedings of the National Academy of Sciences, 83*(20), 7984–7988.

Victor, J. D., Chubb, C., & Conte, M. M. (2005). Interaction of luminance and higher order statistics in texture discrimination. *Vision Research, 45*, 311–328.

Victor, J. D., & Conte, M. M. (1989). Cortical interactions in texture processing: Scale and dynamics. *Visual Neuroscience, 2*(3), 297–313.

Victor, J. D., & Conte, M. M. (1991). Spatial organization of nonlinear interactions in form perception. *Vision Research, 31*, 1457–1488.

Victor, J. D., & Conte, M. M. (2012). Local image statistics: Maximum-entropy constructions and perceptual salience. *Journal of the Optical Society of America A, 29*(7), 1313–1345.

Victor, J. D., Thengone, D. J., Rivi, S. M., & Conte, M. M. (2015). A perceptual space of local image statistics. *Vision Research, 117*, 117–135.

Yu, Y., Schmid, A. M., & Victor, J. D. (2015). Visual processing of informative multipoint correlations arises primarily in v2. *eLife, 4*, e06604.