



## Image segmentation driven by elements of form

Jonathan D. Victor\*, Syed M. Rizvi, Mary M. Conte

Feil Family Brain and Mind Research Institute, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10065, United States



### ARTICLE INFO

No. of Reviewers: 2

#### Keywords:

Visual textures  
Image segmentation  
Curvature  
Closure  
Convexity

### ABSTRACT

While luminance, contrast, orientation, and terminators are well-established features that are extracted in early visual processing and support the parsing of an image into its component regions, the role of more complex features, such as closure and convexity, is less clear. A main barrier in understanding the roles of such features is that manipulating their occurrence typically entails changes in the occurrence of more elementary features as well. To address this problem, we developed a set of synthetic visual textures, constructed by replacing the binary coloring of standard maximum-entropy textures with tokens (tiles) containing curved or angled elements. The tokens were designed so that there were no discontinuities at their edges, and so that changing the correlation structure of the underlying binary texture changed the shapes that were produced. The resulting textures were then used in psychophysical studies, demonstrating that the resulting feature differences sufficed to drive segmentation. However, in contrast to previous findings for lower-level features, sensitivities to increases and decreases of feature occurrence were unequal. Moreover, the texture-segregation response depended on the kind of token (curved vs. angular, filled-in vs. outlined), and not just on the correlation structure. Analysis of this dependence indicated that simple closed contours and convex elements suffice to drive image segmentation, in the absence of changes in lower-level cues.

### 1. Introduction

Visually-guided behavior rests on high-level interpretations of the sensory input – interpretations that require segmenting visual space into meaningful regions, recognizing objects and materials, and estimating their surface properties (Adelson, 2001; Karklin & Lewicki, 2009; Treisman, 1982). These high-level interpretations result from operations carried out on internal representations of the sensory stimulus. While there is broad agreement that internal representations are largely built from local features, there is considerable uncertainty as to what these local features are. Some, such as spatial frequency content and oriented edges, are firmly established (Julesz, 1981). Additionally, it has been suggested that higher-level features, such as extended contours, convexity/concavity, smoothness, and closure, are also important (Barenholtz, Cohen, Feldman, & Singh, 2003; Bertamini & Wagemans, 2013; Hulleman, te Winkel, & Boselie, 2000; Sharan, Liu, Rosenholtz, & Adelson, 2013; Treisman & Gormican, 1988; Wagemans et al., 2012). However, identifying and quantifying the possible contributions of these latter features is more difficult.

Two factors likely contribute to this. First, neurophysiologic studies provide a mechanistic basis for extraction of spatial frequency content and orientation: local linear filters, followed by a simple nonlinearity,

are a reasonable first approximation for many retinal, thalamic, and cortical neurons (Rust, Schwartz, Movshon, & Simoncelli, 2005), and filters are strongly oriented in visual cortex. However, these descriptions are admittedly caricatures, and do not account for the responses of real neurons to naturalistic stimuli (Freeman, Ziemba, Heeger, Simoncelli, & Movshon, 2013; Nirenberg & Pandarinath, 2012; Willmore, Prenger, & Gallant, 2010) or stimuli with complex spatial structure (Freeman et al., 2013; Yu, Schmid, & Victor, 2015). Thus, the potential role of cortical neurons in extracting features that cannot immediately be expressed in terms of orientation and spatial frequency content is less clear.

The second reason for this uncertainty is that features such as closure and convexity are difficult to disentangle from other elementary features, such as orientations and terminators. For example, the visual “pop-out” of an O amidst an array of C’s could be driven by closure (Fig. 11 of Treisman and Gormican (1988)), but these two tokens also differ in the number of line terminators. When this confound is removed, and closure is manipulated in a way that does not alter the number of terminations (Fig. 6 of Julesz (1981)), “pop-out” is markedly reduced if not eliminated.

This confound is part of the general problem that intuitively-defined local features are interdependent. Just as corners require edges,

\* Corresponding author.

E-mail address: [jdvicto@med.cornell.edu](mailto:jdvicto@med.cornell.edu) (J.D. Victor).

<https://doi.org/10.1016/j.visres.2018.12.003>

Received 19 October 2018; Received in revised form 30 November 2018; Accepted 4 December 2018

Available online 01 April 2019

0042-6989/ © 2018 Elsevier Ltd. All rights reserved.

changing a contour from an open C to a closed O affects the number of terminators. A systematic way to address this problem is to work in a stimulus domain that takes these geometry-implied interdependencies into account.

Previously, we had developed such a domain (Victor & Conte, 2012), and showed that it led to a simple picture of how elementary local features (luminance, edge, and corner) interact. In this approach, synthetic visual textures consisting of black and white checks were specified by a set of local spatial correlations, and textures were constructed to be as random as possible given these correlations. We considered all of the local correlations (image statistics) needed to determine the probabilities of  $2 \times 2$  neighborhoods of checks (Victor & Conte, 2012). This resulted in 10 parameters (dimensions) in the space, and allowed for the independent manipulation of mean luminance, horizontal and vertical edges, and corners. In this space, human perceptual thresholds take on a simple form (Victor, Thengone, Rizvi, & Conte, 2015). There are cardinal axes of sensitivity, consistent across observers, each reflecting specific mixtures of these local features. For variations in features that are oblique to these axes, sensitivity is determined by projections onto the cardinal axes. These projections combine in a quadratic fashion, yielding approximately ellipsoidal isodiscrimination contours.

Here we extend this strategy to examine the possible roles of closure and convexity. A “brute-force” extension – in which we consider larger neighborhoods of checks and therefore higher orders of correlation – is evidently impractical. The problem is that the dimensionality of the space rapidly becomes enormous (for example, for correlations in a  $3 \times 3$  neighborhood, there are 400 dimensions), and none of these dimensions correspond to closure or convexity *per se*. Instead, we take an approach that retains the advantages of controlling for confounding effects of lower-level features, and enables us to focus on the features of interest. For example, we would want to change the frequency of closed contours, without changing the number of terminators. To do this, we build on the construction of Victor and Conte (2012), taking the general approach of replacing the black and white checks of the binary textures by tiles that contain sectors of circles or other simple shapes (Figs. 1 and 2). The tiles are designed so that they match along their edges. Thus, no terminators are formed. However, even though there are no terminators, the extent to which the tiled textures contain closed contours depends on the spatial correlations of the starting texture. Moreover, for the main comparisons of interest, the orientation content of all the textures is identical.

Below we report perceptual thresholds on a segmentation task using this stimulus set. We find that the occurrence of small closed contours is a strong cue to segmentation, even when luminance and orientation cues are balanced, and terminators are absent. By manipulating the shapes within the tiles, we also identify an interaction between shape and border ownership, which we interpret as a role for convexity.

## 2. Materials and Methods

Experiments were designed to quantify the extent to which the statistics of specific local visual features supported image segmentation. The stimuli were based on a class of synthetic textures developed in our lab (Victor & Conte, 2012) consisting of arrangements of black and white checks, which we extend here to incorporate confluent micropatterns consisting of curved and angular tokens. The psychophysical paradigm, a four-alternative figure-ground segmentation task, was adapted from the one developed by Chubb, Landy, and Econopouly (2004) and identical to what was used in recent studies (Victor et al., 2015; Victor, Rizvi, & Conte, 2017; Victor, Thengone, & Conte, 2013); we describe it below for the reader’s convenience. Analysis procedures, also identical to procedures previously reported, are likewise summarized below.

The class of stimuli we construct differs not only from classical textures based on isolated micropatterns e.g., Caelli and Julesz (1978),

Julesz (1981), Kingdom, Keeble, and Moulden (1995), Nothdurft, Gallant, and Van Essen (2000), Treisman and Gormican (1988) but also from stimuli constructed as the product of a carrier and an envelope (Hallum, Landy, & Heeger, 2011; Landy & Oruc, 2002; Sutter, Sperling, & Chubb, 1995; Westrick, Henry, & Landy, 2013), a strategy that is especially suitable for analyzing models for spatial vision based on a “filter-rectify-filter” (FRF) cascade. These cascades extract features defined by contrast modulation, because the rectification stage acts as a demodulator, converting both positive or negative fluctuations in contrast to an unsigned quantity that can then be extracted. In this approach, features defined by contrast modulation are known as “second-order” features, because they are extracted by the second linear filter. Here, we don’t restrict ourselves to this class of features. We also use the term “second-order” in a more restricted, mathematical sense – statistical properties of images that can be gleaned by simultaneous analysis of pairs of points.

### 2.1. Texture stimuli

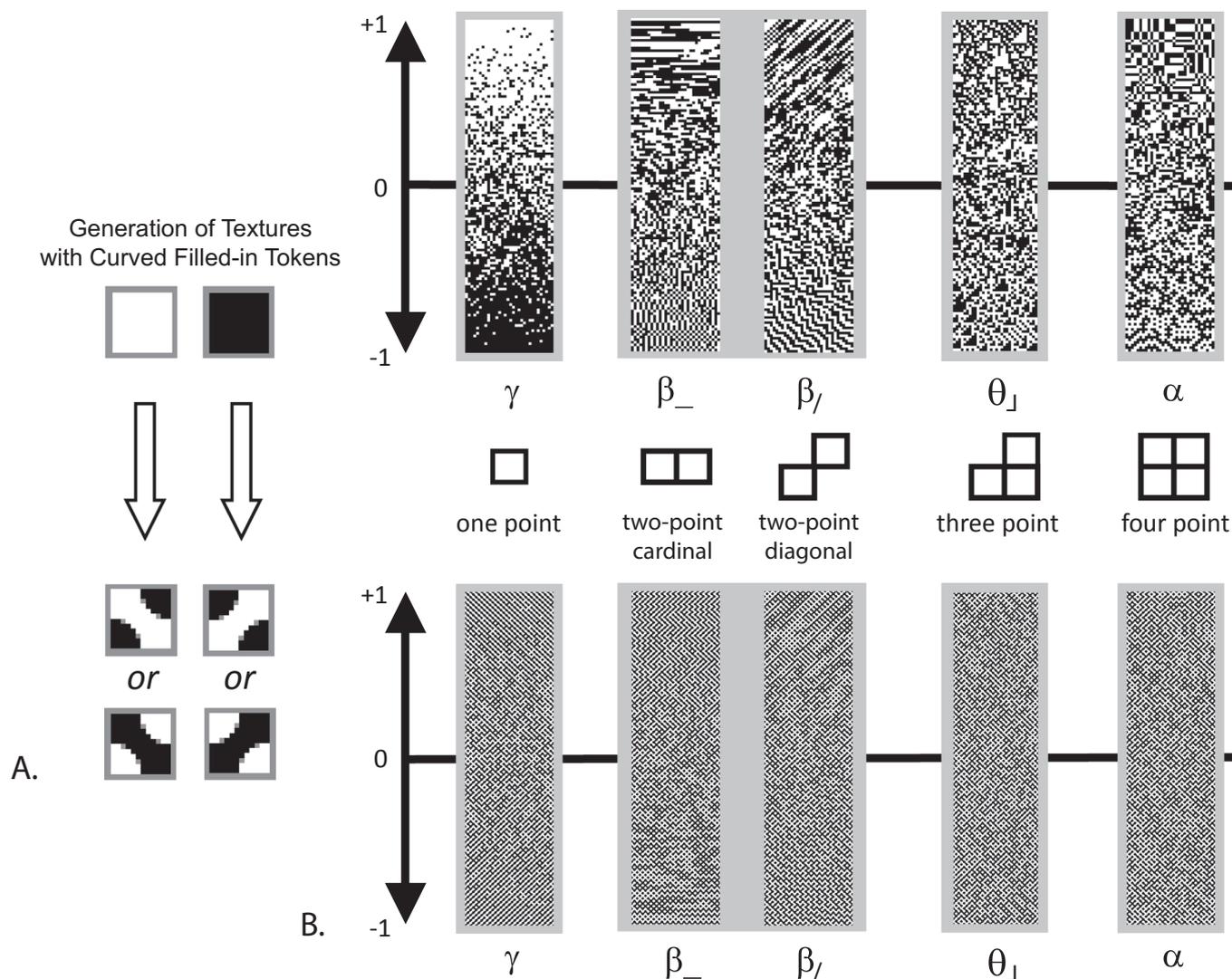
Stimuli were constructed in two stages: first, creation of underlying binary textures of white and black checks, and second, replacing the checks with tokens. If the second stage is omitted – i.e., if the tokens consist of white and black checks – the resulting textures are identical to those used in previous studies (Victor & Conte, 2012; Victor, Rizvi, et al., 2017; Victor et al., 2013; Victor et al., 2015) and referred to here as the “standard” textures. For the main experiments, the tokens consist of sectors of circles, triangles, or lines. We describe the token and token-replacement procedure first, and then summarize how the underlying binary textures are generated.

### 2.2. Tokens

Fig. 1 shows the procedure for replacing the standard black and white checks by tokens. For the first series of experiments, the tokens contained curved contours and filled-in regions (as shown in Fig. 1A). White checks were replaced by tokens with a solid region running from upper left to lower right; black checks were replaced by tokens with a solid region running from upper right to lower left. The polarity of each check was chosen so that there were no discontinuities at the edges of the checks. This produced texture patterns with curved contours running continuously across the check boundaries, designated the “curved filled-in” texture. Depending on the binary values in the underlying texture, the resulting regions were convex or concave, and could range from simple circles to snake-like shapes running through the entire texture (see examples in Fig. 1B, and column 2 of Fig. 2). Note that the tokens contain an equal amount of white and black, and that it is always possible to place the tokens so that there are no discontinuities at the check boundaries.

We also used several variations on the tokens shown in Fig. 1. One variation was to replace the interiors of the regions of the curved filled-in tokens by white and to replace the boundaries between them by black, resulting in “curved outline” tokens and textures (column 3 of Fig. 2). The contour shapes are identical, but when the filled-in regions are replaced by outlines, there is a reduced sense of border ownership. A second variation was to replace the curved contours by straight line segments; this changed the rounded contours of the resulting texture to short line segments that met at right angles (“angular filled-in” and “angular outline” textures). These variations are shown in columns 4 and 5 of Fig. 2.

We note that this strategy of replacing an underlying texture of checks by tokens was initially used by Julesz, Gilbert, and Victor (1978) (their Fig. 4), but with a different set of tokens: squares containing line segments at  $+45$  degrees or  $-45$  degrees. With these tokens, changes in the statistics of the underlying texture results in changing the number of terminators and T-junctions – a possible confound that is avoided with the current token set.



**Fig. 1.** Generation of a texture with curved filled-in tokens. A. Transformation from standard binary textures to curved filled-in token textures. In a standard binary texture, each check is assigned to either black or white. To create the corresponding token texture, each white check is replaced by a square containing a curved filled-in token; white checks are replaced by tokens in which the connected component runs from upper left to lower right; black checks are replaced by tokens in which the connected component runs from upper right to lower left. The components of the tokens are colored in either of two polarities, so that there are no discontinuities at the edges of the checks. Note that the black and the white regions in each check have equal area. B. The stimulus gamuts. The top portion shows selected coordinates of the space of binary textures:  $\gamma$ , the difference between the fraction of white and black checks, is a first-order statistic; the other coordinates correspond to higher-order statistics ( $\beta_-$ , second-order correlations along the horizontal axis;  $\beta_\gamma$ : second-order correlations along a diagonal axis;  $\theta_\gamma$ , a correlation among triplets of checks; and  $\alpha$ , a correlation among the four checks of a  $2 \times 2$  neighborhood). Each strip shows the textures generated by varying each coordinate through its allowable range (-1 to + 1), with 0 corresponding to randomness. Bottom portion: Results of replacing checks by the tokens shown in A. Top portion of Panel B adapted from Fig. 1 of (Victor et al., 2013), with permission of the copyright holder, The Association for Research in Vision and Ophthalmology.

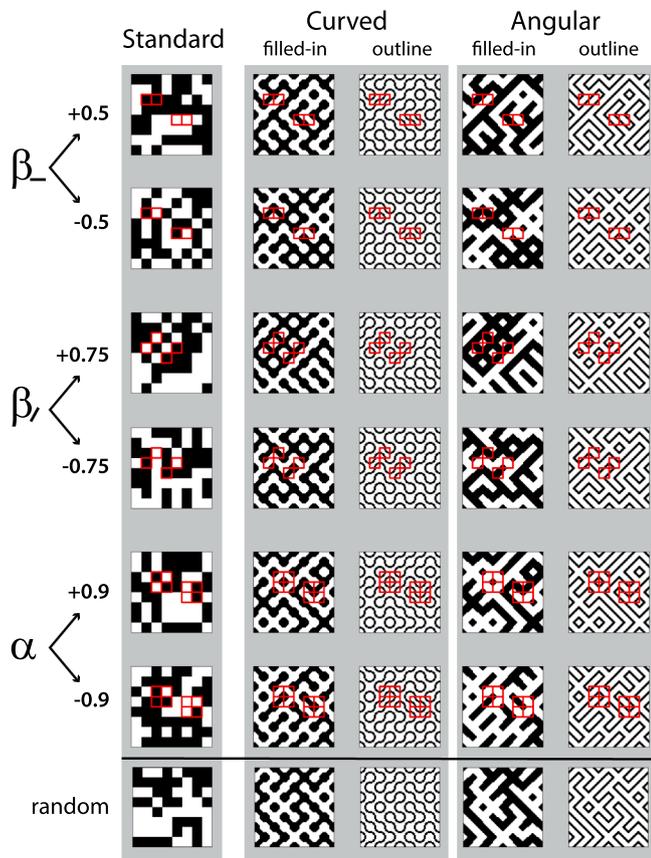
### 2.3. Underlying textures

For the underlying binary textures, we used the space of maximum-entropy textures developed in Victor and Conte (2012). The rationale for this choice is that it provided precise control over the statistical cues that are available for image segmentation: stimuli are specified by their local correlations, and the textures are constructed to be as random as possible given these correlations. Further background and rationale can be found in other publications that use this domain (Hermundstad et al., 2014; Victor, Rizvi, et al., 2017; Victor et al., 2013; Victor et al., 2015); the main points are reproduced here.

As we are focusing on local correlations (i.e., correlations within  $2 \times 2$  neighborhoods), stimuli are specified by the probabilities of occurrence of each of the ways that the four checks in these neighborhoods can be assigned to binary values. Although 16 such assignments are possible ( $16 = 2^{2 \times 2}$ ), there are only 10 degrees of freedom (Victor &

Conte, 2012): these 16 probabilities must sum to 1, and other constraints arise to ensure internal consistency: for example, computing the  $1 \times 2$  block probabilities from the top half of  $2 \times 2$  blocks must yield the same result as computing them from the bottom half. As in the previous studies, we organize these 10 degrees of freedom for the stimuli into groups, corresponding to first-, second-, third-, and fourth-order correlations. Each coordinate ranges from -1 to +1; a value of 0 indicates no correlation. The origin of the space (a value of 0 for each coordinate) is a completely random binary image.

Fig. 1B illustrates selected coordinates, both in the context of the standard texture (top) and a texture in which the checks have been replaced by curved filled-in tokens (bottom). The coordinate  $\gamma$  indicates the first-order correlation, namely, the difference between the probability of the two binary assignments in the underlying texture.  $\gamma = +1$  means that all checks have one assignment (white for the standard texture, contour from upper left to lower right in the curved filled-in



**Fig. 2.** Variations of the token textures. Each row shows an  $8 \times 8$  region of an underlying texture constructed with  $\beta_- = \pm 0.5$  (rows 1 and 2),  $\beta = \pm 0.75$  (rows 3 and 4), and  $\alpha = \pm 0.9$  (rows 5 and 6) for an underlying texture constructed with black and white checks, and for each of the four token types studied. The outlined regions indicate pairs or quadruples of checks that contribute to the statistical structure. For  $\beta_- > 0$  (row 1), adjacent token pairs have contours of similar orientation, and cannot form portions of circles or squares. For  $\beta_- < 0$  (row 2), adjacent token pairs have contours of dissimilar orientations, and can form portions of circles or squares. For  $\beta > 0$  (row 3), tokens that share a corner have contours of similar orientation, and also can form portions of circles or squares. For  $\beta < 0$  (row 4), tokens that share a corner have contours of dissimilar orientation, and cannot form portions of circles or squares. For  $\alpha > 0$  (row 5), adjacent tokens are unconstrained, and  $2 \times 2$  neighborhoods can form circles or squares. For  $\alpha < 0$  (row 6), adjacent tokens are also unconstrained, but  $2 \times 2$  neighborhoods cannot form circles or squares. The bottom row shows a region of a random texture, constructed with all image statistics equal to zero, and its rendering with the four token types.

token texture);  $\gamma = -1$  means that all checks have the opposite assignment, and  $\gamma = 0$  means that both are equally likely. Once the standard checks have been replaced by the curved filled-in tokens,  $\gamma$  biases the orientations of the curved boundaries, but all textures have an equal amount of white and black.

There are four coordinates, denoted  $\beta_-$ ,  $\beta$ ,  $\beta_+$ , and  $\beta_\gamma$ , devoted to second-order correlations. The value of each coordinate is the difference between the probability that two neighboring checks (in the direction indicated by the subscript) have the same assignment and the probability that they do not. For example,  $\beta_- = +1$  means that horizontally-adjacent checks have the same assignment. In the standard texture (top of strip for  $\beta_-$  in upper part of Fig. 1B), this produces rows that contain only a single color; in the curved filled-in token texture (top of strip for  $\beta_-$  in lower part of Fig. 1B), this produces rows that are either a uniform region of slanting from upper left to lower right, or from upper right to lower left – thus forming a zig-zag pattern. In contrast,  $\beta_- = -1$  means that horizontally-adjacent checks always have the opposite assignment (bottom of strip for  $\beta_-$  in upper part of

Fig. 1B). In the standard texture, this produces rows of alternating black and white checks. In the curved filled-in token texture, it produces regions that contain circles and horizontal waves (bottom of strip for  $\beta_-$  in lower part of Fig. 1B). While the effects of  $\beta_-$  and  $\beta_+$  (the “cardinal”  $\beta_-$ ’s) are similar, the effects of  $\beta_+$  and  $\beta_\gamma$  (the “diagonal”  $\beta_-$ ’s) are quite distinct. For the standard texture, values of  $\beta_\pm$  near  $\pm 1$  generate correlations along the diagonal; for the curved filled-in token texture, they generate regions of dumbbell-shaped regions ( $\beta_+$  near +1) and peninsulas ( $\beta_+$  near -1). Thus, substitution of tokens for uniform checks serves to dissociate shape from correlation *per se*. The influences of the second-order statistics can also be seen in Fig. 2, which shows magnified views of smaller portions of the textures. Note that some manipulations increase the frequency of circular closed contours ( $\beta_- < 0$  and  $\beta_+ > 0$ ), while others decrease their frequency ( $\beta_- > 0$  and  $\beta_+ < 0$ ), a point that we will quantify below (see Eq. (3)).

Third-order correlations are quantified by four coordinates,  $\theta_\downarrow$ ,  $\theta_\uparrow$ ,  $\theta_\leftarrow$ , and  $\theta_\rightarrow$ , corresponding to the four possible orientations of an L-shaped region containing three adjacent checks. In the standard texture, a value of +1 means that every such region contains an odd number of white checks, while -1 means that every such region contains an even number of white checks (and an odd number of black checks). For coordinate values near these extremes, the resulting texture has prominent triangular shaped regions (white for  $\theta$  near +1, black for  $\theta$  near -1). However, when the checks are replaced by curved filled-in tokens, the difference between  $\theta = \pm 1$  and  $\theta = 0$  is difficult to discern visually – a point that we confirmed with formal psychophysical measures.

Finally, a single fourth-order coordinate,  $\alpha$ , quantifies the four-point correlations between points in a  $2 \times 2$  region. For  $\alpha = +1$ , all regions of the standard texture contain an even number of white checks, producing rectangular blobs (top of strip for  $\alpha$  in upper part of Fig. 1B). The corresponding curved filled-in token texture (top of strip for  $\alpha$  in lower part of Fig. 1B) contains circles and diamond-shaped closed contours. For  $\alpha = -1$ , all  $2 \times 2$  regions contain an odd number of white checks (bottom of strip for  $\alpha$  in upper part of Fig. 1B) and peninsula- and dumbbell-shaped regions for the curved filled-in token texture. (bottom of strip for  $\alpha$  in lower part of Fig. 1B). Magnified views of the effects of  $\alpha$  are shown in Fig. 2. Circular closed contours are prominent for  $\alpha$  near +1 and rare for  $\alpha$  near -1, as will be quantified below (Eq. (3)).

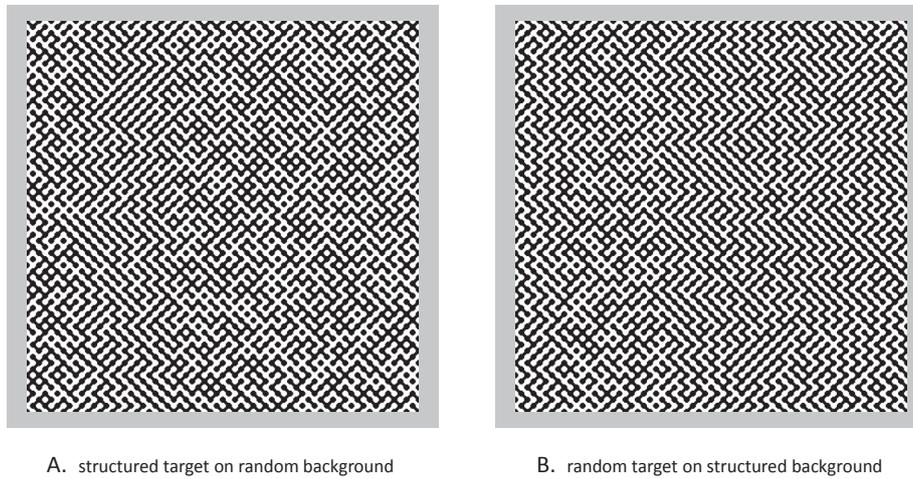
#### 2.4. Segmentation task

To quantify segmentation, we used the paradigm introduced by Chubb and coworkers (Chubb et al., 2004) initially developed for synthetic textures without spatial correlations, but readily applicable to textures with correlated spatial structure (Victor, Chubb, & Conte, 2005; Victor & Conte, 2012; Victor, Rizvi, et al., 2017; Victor et al., 2013; Victor et al., 2015) and tokens, as summarized here.

Stimuli consisted of  $64 \times 64$  arrays, each containing an embedded  $16 \times 64$  rectangular target whose outer edge was 8 checks from one side of the array. The subject’s task was to indicate the target position via a button-press on a four-button response box.

Checks of the target and background regions were filled with tokens either randomly, or according to a specified nonzero value of one or more of the texture parameters described above (Fig. 3). Half of the trials had a structured target on a random background (Fig. 3A); half had a random target on a structured background (Fig. 3B). This was done to ensure that the subject performed the task by identifying a texture boundary, rather than a texture gradient (Wolfson & Landy, 1998) – as the latter strategy would not yield a fraction correct greater than 0.5. Our previous work has shown no consistent threshold difference between these conditions. Therefore, analyses are based on pooling across this randomization.

In most experiments (“single-axis” blocks), we measured sensitivity to variation of a single texture parameter. The values of the statistics were set at  $\pm 0.3c_{max}$ ,  $\pm 0.6c_{max}$ ,  $\pm 0.8c_{max}$ , and  $\pm 0.95c_{max}$ , where  $c_{max} = 1$  for all of the token textures and for the standard textures along



**Fig. 3.** Examples of two typical stimuli constructed with the curved filled-in tokens. The stimulus is an array of  $64 \times 64$  tokens, and the target is a vertical strip of  $16 \times 64$  tokens, beginning 8 tokens away from the left-hand edge. In A, the target is structured ( $\beta_- = 0.72$ ) and the background is random; in B, the target is random and the background is structured (also with  $\beta_- = 0.72$ ).

the  $\theta_j$ - and  $\alpha$ -axes,  $c_{max} = 0.25$  for the standard texture along the  $\gamma$ -axis,  $c_{max} = 0.5$  for the standard texture along the  $\beta_-$ -axis, and  $c_{max} = 0.75$  for the standard texture along the  $\beta_j$ -axis. (Values of  $c_{max} < 1$  were used along the axes in which we anticipated lower thresholds based on previous data.)

In other experiments (“coordinate plane” blocks), we measured sensitivity to joint variation of two texture parameters:  $\{\beta_-, \beta_j\}$  or  $\{\beta_j, \beta_j\}$ . Here, values of the statistics were set at  $\pm 0.2c_{card}$ ,  $\pm 0.4c_{card}$ ,  $\pm 0.6c_{card}$ ,  $\pm 0.8c_{card}$ , and  $\pm c_{card}$  along the cardinal axes (where only one statistic had a nonzero value), where  $c_{card} = 0.45$  for standard textures in the  $(\beta_-, \beta_j)$ -plane,  $c_{card} = 0.75$  for standard textures in the  $(\beta_j, \beta_j)$ -plane, and  $c_{card} = 1$  for all of the token textures. For exploration of these planes in oblique directions, statistics were set at  $(\pm 0.7c_{obl}, \pm 0.7c_{obl})$  and  $(\pm c_{obl}, \pm c_{obl})$ , where  $c_{obl} = 0.24$  for standard textures in the  $(\beta_-, \beta_j)$ -plane,  $c_{obl} = 0.4$  for standard textures in the  $(\beta_j, \beta_j)$ -plane, and  $c_{obl} = 1$  for all of the token textures. For the standard tokens, these choices matched the conditions used in Victor et al. (2015).

In all cases, the texture coordinates not specified above were determined by the maximum-entropy construction of Victor and Conte (2012) (see its Table 2).

## 2.5. Procedure

As in previous studies in our lab (Victor et al., 2015), we used a Cambridge Research ViSaGe system, running custom Delphi software, to produce the stimuli and collect responses. The monitor was a LCD monitor (mean luminance of  $23 \text{ cd/m}^2$ , refresh rate 60 Hz). Stimulus duration was 120 ms, followed by a 500-ms random mask consisting of checks that were half the size of the stimulus checks. The display size was  $15 \times 15 \text{ deg}$  ( $64 \times 64$  checks, 14.8 min each); viewing distance was 103 cm, and contrast was 1.

The experimental procedure was identical to that of Victor et al. (2015), and is summarized here. All trials were self-paced, triggered by a button-press. 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 with a button-press on a four-button response box. They were asked to fixate centrally and not attempt to scan the stimulus. Inexperienced subjects received training of approximately two hours to become accustomed to the brief stimulus presentation time and to learn to maintain central fixation without scanning. During this training, but not during data collection, subjects received auditory feedback for incorrect responses.

Blocks of trials consisted of stimuli constructed with a single type of

token or with uniform checks. In a single-axis block, one of the parameters  $\gamma$ ,  $\beta_+$ ,  $\beta_j$ ,  $\theta_j$ , or  $\alpha$  was set to a nonzero value, as specified above: these blocks had 320 trials: the above 5 parameters, 8 values for each parameter ( $\pm 0.3c_{max}$ ,  $\pm 0.6c_{max}$ ,  $\pm 0.8c_{max}$ , and  $\pm 0.95c_{max}$ , with  $c_{max}$  specified above), 4 target positions, each presented either with target structured or background structured. In coordinate-plane blocks, two parameters, drawn either from the set  $\{\beta_-, \beta_j\}$  or the set  $\{\beta_j, \beta_j\}$ , were jointly varied. These blocks had 288 trials: 36 coordinate pairs (20 on-axis parameter values and 8 off-axis values each repeated twice), in 4 target positions, each presented either with target structured or background structured. 15 blocks of each condition were run, with different stimulus examples in each block. Trial order was randomized within each block, and block order was counterbalanced across subjects.

## 2.6. Analysis

For each texture type (the standard texture and the four kinds of tokens), we quantified the threshold for segregation by the procedure of Victor et al. (2005, 2013, 2015), as summarized here. First, along each ray  $r$  emanating from the origin of the texture space (i.e., for each series of increasing correlation strengths), we determined, via maximum-likelihood, the best-fitting Weibull function to the observed fraction correct (FC):

$$FC(x) = \frac{1}{4} + \frac{3}{4}(1 - 2^{-(x/a_r)^{b_r}}), \quad (1)$$

where  $x$  is the distance between the test and reference point,  $a_r$  is the fitted threshold (i.e., the value of  $x$  at which  $FC = 0.625$ , halfway between chance (0.25), and perfect (1.0)), and  $b_r$  is the Weibull shape parameter. The distance  $x$  is the Euclidean distance from the origin of the space: in a single-axis experiment, it is the absolute value of the texture coordinate; in a coordinate plane experiment, it is given by  $x = \sqrt{c_y^2 + c_z^2}$ , where  $c_y$  and  $c_z$  are the values of the two coordinates, drawn from  $\{\beta_-, \beta_j\}$  or  $\{\beta_j, \beta_j\}$ . The shape parameter  $b_r$  typically had similar values across rays – for the standard textures (as previously reported Victor et al. (2015)), confidence intervals overlapped heavily and usually included the range 2.2–2.7; a similar behavior was seen for the textures constructed with tokens. Thus, to focus on thresholds, we 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. 95% confidence intervals were determined via 1000-sample bootstraps. This procedure reduced the number of free parameters in the fits without altering their quality, and resulted in smaller confidence intervals for the threshold. Note that this

procedure could result in a measured threshold that was greater than 1, if performance was above chance but never reached a FC of 0.625. We defined sensitivity as  $1/\text{threshold}$ , with corresponding confidence intervals.

Across-subject averages of sensitivities or thresholds are computed as the geometric means, and statistics (standard errors, *t*-tests, ANOVA) are computed on the logarithms of the raw values.

## 2.7. Subjects

Studies were conducted in 6 normal subjects (2 male, 4 female), ages 21 to 55. Three of the subjects (MC, SR, KP) were experienced psychophysical observers; SP had approximately 5 h experience (as part of Victor et al. (2015)), and the other subjects (NM, WC) had no previous experience. MC and SR are authors; the other four observers were naïve to the purposes of the experiment. All subjects had visual acuities (corrected if necessary) of 20/20 or better.

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.

## 3. Results

We carried out three related experiments to probe visual sensitivity to elements of form. All experiments consisted of measuring segmentation thresholds for textures defined by their local image statistics. As detailed in Methods, the textures are generated by specifying correlations between nearest-neighbor checks. To probe form, the checks are filled by tokens consisting of simple micropatterns, which are designed so that changing the correlations between them modulates the occurrence of extended contours and simple shapes, but does not change the occurrence of features such as luminance, contrast, or orientation content.

The first experiment considers just one kind of token, and explores the effects of correlations involving one, two, three, and four nearest-neighbor checks. As second-order interactions are found to be the most revealing, we focus on them for the remaining experiments. The second set of experiments compares different kinds of tokens, which probe the relevance of smooth vs. angular contours, and filled-in vs. outlined shapes. The third set of experiments returns to the original set of curved tokens, and examines how different kinds of two-check correlations interact.

### 3.1. Experiment 1

This experiment surveyed segmentation driven by textures in which a single kind of correlation governed the placement of curved filled-in tokens. Fig. 4 provides a detailed view of typical data in two subjects (SP, KP), and illustrates the main findings. Each plot is a psychometric function for targets defined by correlations between pairs of checks that either share an edge ( $\beta_{\perp}$ ) or a corner ( $\beta_{\angle}$ ). For standard checks, performance is nearly identical for positive correlations and for negative correlations. When the checks are replaced by curved tokens, thresholds increase. But also, in contrast to the polarity-independence seen for standard textures, performance also depends on whether the correlation is positive or negative. This polarity-dependence is most evident for  $\beta_{\angle}$ : performance is better (lower thresholds) for positive correlations than for negative correlations.

Fig. 5 summarizes the performance data for all subjects and correlation types, taking the threshold as the correlation strength at which the fraction correct is 0.625 (see Methods). Each panel corresponds to a particular type of correlation, and compares thresholds for textures containing curved tokens (right) with thresholds for standard textures (left). Thresholds for standard checks for four of the six subjects were previously reported in Fig. 3 of Victor et al. (2015), and are taken from

that Figure. Thresholds for MC and SR are new determinations made along with the current measurements of thresholds for curved tokens.

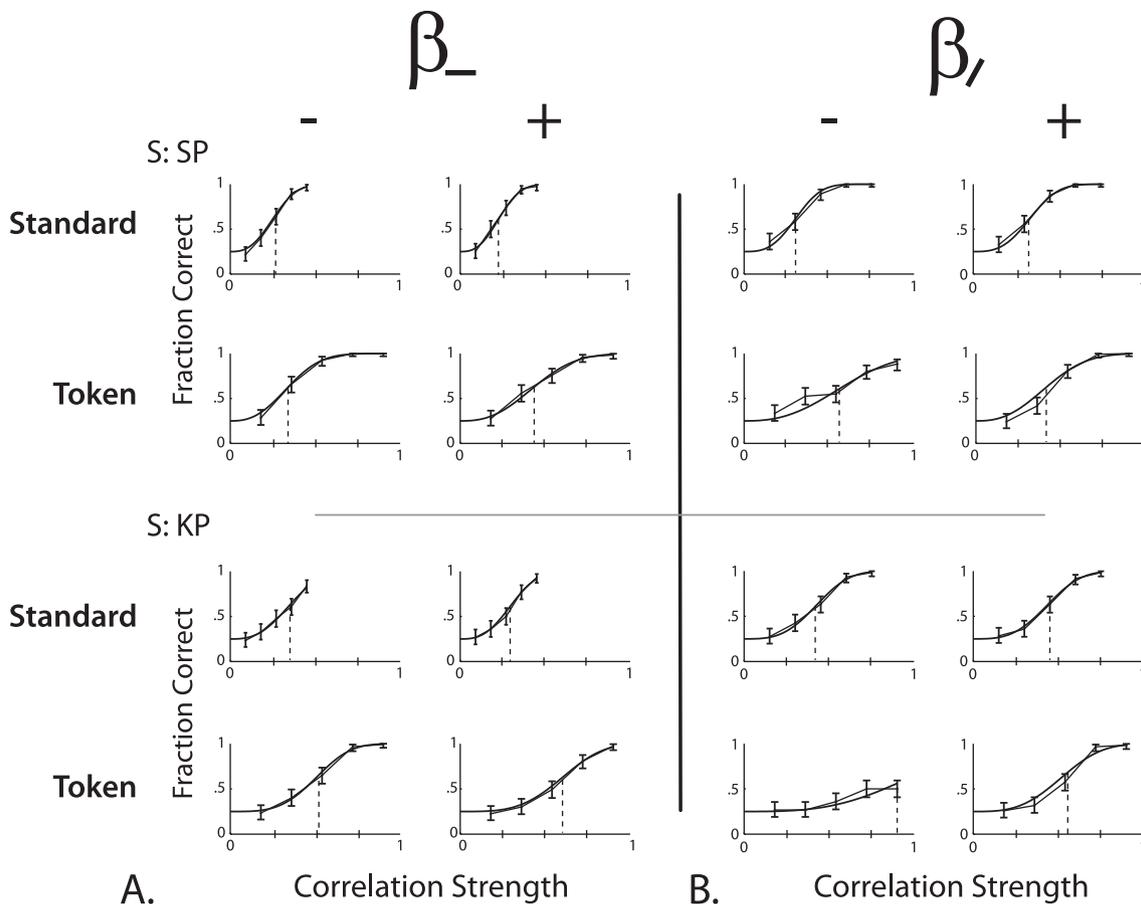
The first-order statistic  $\gamma$  (Fig. 5 upper panel) controls the fraction of checks of each type (one check type has probability  $(1 + \gamma)/2$ ; the other has  $(1 - \gamma)/2$ ). Consistent with previous results, Victor et al. (2015), thresholds for the standard texture were  $0.183 \pm 27\%$  (geometric mean  $\pm$  fractional SD). There is a trend towards a lower threshold for negative values of  $\gamma$  (i.e., a bias towards black checks); thresholds were  $0.177 \pm 26\%$  for black checks and  $0.189 \pm 31\%$  for white checks; two-tailed paired *t*-test:  $p < 0.08$ . A similar but statistically significant difference was previously reported in a larger group of subjects (Victor et al., 2015). When the checks are replaced by curved tokens, thresholds rise to  $0.498 \pm 13\%$ . There is no detectable asymmetry between negative and positive deviations of  $\gamma$  ( $0.497 \pm 13\%$  for negative values,  $0.499 \pm 14\%$  for positive values; two-tailed paired *t*-test:  $p > 0.5$ ). This lack of dependence on polarity is expected since the two polarities of  $\gamma$  correspond to leftward-up-sloping and rightward-up-sloping orientations of the dominant contours of the tokens.

Second-order statistics, consistent with the findings detailed in Fig. 4, show a contrasting behavior (Fig. 5 middle panels). Thresholds are polarity-independent for the standard texture ( $\beta_{\perp}$ ,  $0.295 \pm 25\%$ ;  $\beta_{\angle}$ ,  $0.431 \pm 20\%$ ; two-tailed paired *t*-test comparing thresholds for negative and positive polarities,  $p > 0.5$  in both cases) consistent with Victor et al. (2015). In contrast, thresholds are strongly polarity-dependent for checks consisting of curved tokens. This polarity-dependence in turn depended on whether the correlation involved checks that shared an edge (the horizontal second-order correlation  $\beta_{\perp}$ ), or checks that shared a corner (the diagonal second-order correlation  $\beta_{\angle}$ ). In the former case ( $\beta_{\perp}$ ), thresholds were lower for negative correlations:  $0.500 \pm 18\%$  compared to  $0.571 \pm 15\%$ ; two-tailed paired *t*-test:  $p < 0.005$ . In the latter case ( $\beta_{\angle}$ ), the findings were opposite and the difference was larger:  $0.917 \pm 22\%$  for negative correlations, and  $0.560 \pm 13\%$  for positive correlations; two-tailed paired *t*-test,  $p < 0.001$ . Note that although the sign of the correlation that improved performance depended on whether the correlation was horizontal ( $\beta_{\perp}$ ) or diagonal ( $\beta_{\angle}$ ), in both cases the lower-threshold condition was characterized by a greater number of circles (Fig. 2 rows 2 and 3).

For third-order statistic  $\theta_{\perp}$ , most thresholds were unmeasurably high for curved tokens and performance was close to chance even for the most saturated conditions ( $\theta_{\perp} = \pm 0.95$ ). There was only one instance in one subject ( $\theta_{\perp} < 0$ , MC) for which the threshold was under 1.8. For standard checks, consistent with (Victor et al., 2015), thresholds were  $0.731 \pm 20\%$  (5 subjects, NM not measured). These results are not shown.

For the fourth-order statistic  $\alpha$ , performance for the curved tokens was better than chance, but more variable across subjects than for first- and second-order statistics. Four subjects had higher thresholds for  $\alpha < 0$  than for  $\alpha > 0$ ; in three of these subjects, the difference was approximately 20%; in the fourth, the threshold for  $\alpha < 0$  was unmeasurably high as performance was at chance for  $\alpha = -0.95$ . One subject (WC) had unmeasurably high thresholds for both positive and negative values of  $\alpha$ . The sixth subject (NM) had unmeasurably high thresholds for  $\alpha > 0$  and a high but measurable threshold of 1.89 for  $\alpha < 0$ . As in the results for lower-order statistics, findings for standard checks were consistent with Victor et al. (2015):  $0.771 \pm 43\%$  for negative correlations, and  $0.571 \pm 23\%$  for positive correlations; two-tailed paired *t*-test:  $p < 0.02$ .

In sum, in contrast to our findings for segmentation driven by differences in the correlation between nearest-neighbor black and white checks, thresholds for segmentation driven by correlations between curved tokens depended strongly on the polarity of the correlation (Fig. 4, and middle panels of Fig. 5). Moreover, this dependence was one of lower thresholds for negative than for positive correlations when the checks shared a common edge (data for  $\beta_{\perp}$  in Fig. 5), but higher thresholds for negative than for positive correlations when the checks shared a common corner (data for  $\beta_{\angle}$  in Figs. 4 and 5). The data for the



**Fig. 4.** Psychometric functions for standard checks and curved filled-in tokens for cardinal second-order correlation  $\beta_-$  (panel A) and diagonal second-order correlation  $\beta_$  (panel B). In each case, psychometric functions are shown for positive and negative correlations; chance performance is 0.25. Error bars are 95% confidence limits. The smooth curve, which sometimes obscures the data points, is the Weibull function fit to the data, and the vertical dashed lines indicate the thresholds, where the fraction correct reaches 0.625. Subjects SP and KP.

fourth-order correlation (data for  $\alpha$  in Fig. 5) showed a trend that was similar to the findings for  $\beta_$ .

To begin to interpret these findings, we note that for the token textures, manipulation of local images statistics affects the occurrence of several kinds of features, including extended oriented contours, convex contours, and circles. Positive values of  $\beta_-$  increase the likelihood that adjacent tokens contain contours with similar orientations, and therefore result in an extended contour. Conversely, negative values of  $\beta_-$  increase the likelihood that adjacent tokens form a local convexity. The occurrence of these features is not modulated either by  $\beta_$  or  $\alpha$ , since those parameters do not affect the correlations between two tokens that share an edge.

The occurrence of circles, on the other hand, depends on all three parameters ( $\beta_-$ ,  $\beta_$ , and  $\alpha$ ). To form a circle, tokens that share an edge must be of opposite type, and tokens that share a corner must be of the same type. Thus, because  $\beta_-$  controls correlations across common edges,  $\beta_- < 0$  increases the probability of configurations that can form a circle, while  $\beta_- > 0$  decreases their probability. Similarly, because  $\beta_$  controls correlations across checks along a diagonal that share a common corner,  $\beta_ < 0$  decreases the probability of configurations that can form a circle, while  $\beta_ > 0$  increases their probability. Finally, since a circle consists of two tokens of each type, circles will be more common when  $\alpha > 0$  and less common when  $\alpha < 0$ .

These considerations suggest a preliminary account of the data of Fig. 5: circles lower the threshold for segmentation, since thresholds are lower for  $\beta_- < 0$ ,  $\beta_ > 0$ , and  $\alpha > 0$ .

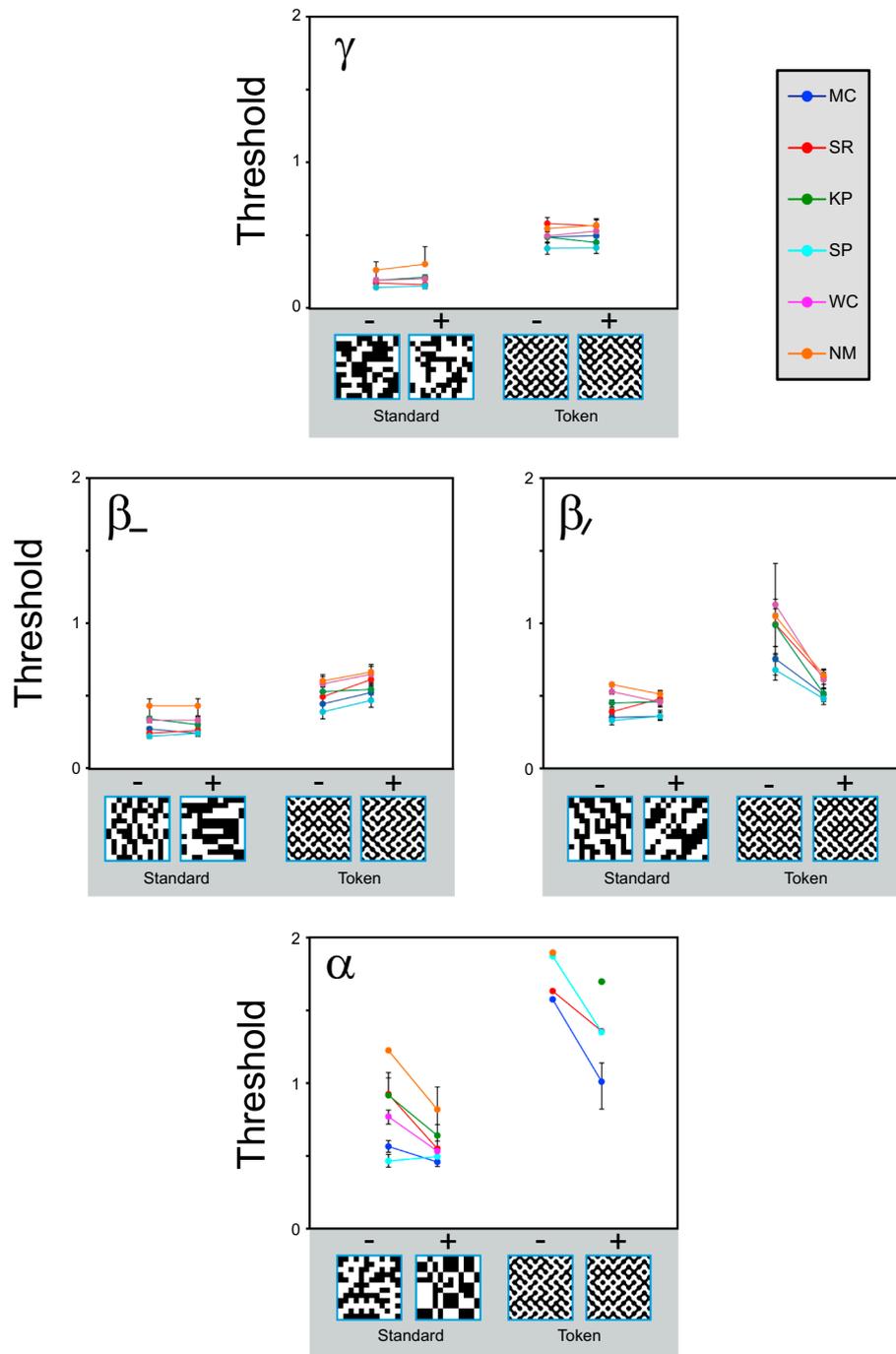
Note that stimuli that are created by varying  $\beta_-$ ,  $\beta_$ , or  $\alpha$  are equated for luminance, contrast, and orientation content – as both target and

background contain the same fraction of each kind of token. Further, there are no “terminators” or junctions in any stimuli, as the edges of adjacent tokens always match to form a continuous region.

### 3.2. Experiment 2

Motivated by the idea that the presence of circles appeared to drive segmentation thresholds, we examined performance for variations on the tokens used in Experiment 1. In the first variation, we replaced the white and black regions of each token by curved segments that ran along their border; we designate these as “curved outline” tokens. In the textures they generate, contours had a similar shape to those of Experiment 1, but, because the regions defined by the contours were no longer filled in, they were not as likely to be seen as objects. In the second variation, we replaced the curved contours by straight line segments; we designate these as “angular filled-in” and “angular outline” tokens. This replaces the undulating contours of the original textures by straight lines, replaces semicircular segments by right angles, and replaces circles by diamonds. We focused on second-order interactions because they were readily measurable and consistent across subjects in Experiment 1.

Results are shown in Fig. 6. Using tokens consisting of outlines rather than filled-in regions increased the thresholds for both cardinal ( $\beta_-$ ) and diagonal ( $\beta_$ ) second-order statistics. For  $\beta_-$ , this modification resulted in an inversion of the asymmetry seen for filled-in tokens: thresholds for positive correlations are lower than for negative correlations. For  $\beta_$ , thresholds for textures constructed with outlines are higher and showed the same direction of asymmetry as filled-in



**Fig. 5.** Threshold data for standard checks and curved filled-in tokens. Stimuli are illustrated by samples of  $12 \times 12$  checks beneath each set of data points. Upper panel: the first-order statistic  $\gamma$ ; middle panels: cardinal and diagonal second-order correlations  $\beta_-$  and  $\beta_+$ ; lower panel: fourth-order correlation  $\alpha$ . Error bars indicate 95% confidence limits by bootstrap. In the lower panel, missing data points correspond to thresholds that were unmeasurably high, and missing error bars indicate that more than 5% of the bootstrap samples led to thresholds that were unmeasurably high. 6 subjects. Example texture patches under the abscissae have correlation strength  $\pm 0.25$  for  $\gamma$ ,  $\pm 0.5$  for  $\beta_-$ ,  $\pm 0.75$  for  $\beta_+$ , and  $\pm 0.9$  for  $\alpha$ .

textures: lower thresholds for positive correlations in either case. While the finding for  $\beta_+$  is consistent with the simple hypothesis that the presence of circles is a key feature in determining segmentation thresholds, the inversion of the influence of  $\beta_-$  when filled-in tokens are replaced by outlines indicates additional factors must also be relevant. We return to this point below.

Replacing the curved contours by straight line segments – which replaces semicircular contours by right angles, and circles by diamonds – had a simpler effect: it reduced thresholds in all conditions, and reduced but did not change the sign of any asymmetries between positive and

negative correlations. In other words, for filled-in tokens, whether curved and angular, thresholds are lower under conditions in which small closed shapes (circles or diamonds) are formed. When either kind of filled-in token is replaced by the corresponding outline token, thresholds rise, and the effect of  $\beta_-$  switches, facilitating segmentation when  $\beta_- > 0$ .

Table 1 shows the results of an ANOVA applied to the data of Fig. 6, and demonstrates that the interactions described above are substantial, statistically significant, and consistent across subjects. The five factors consisted of subject (Sbj) and four stimulus-related factors: BeO: orientation of second-order correlation ( $\beta_-$  vs.  $\beta_+$ ); Pol: correlation

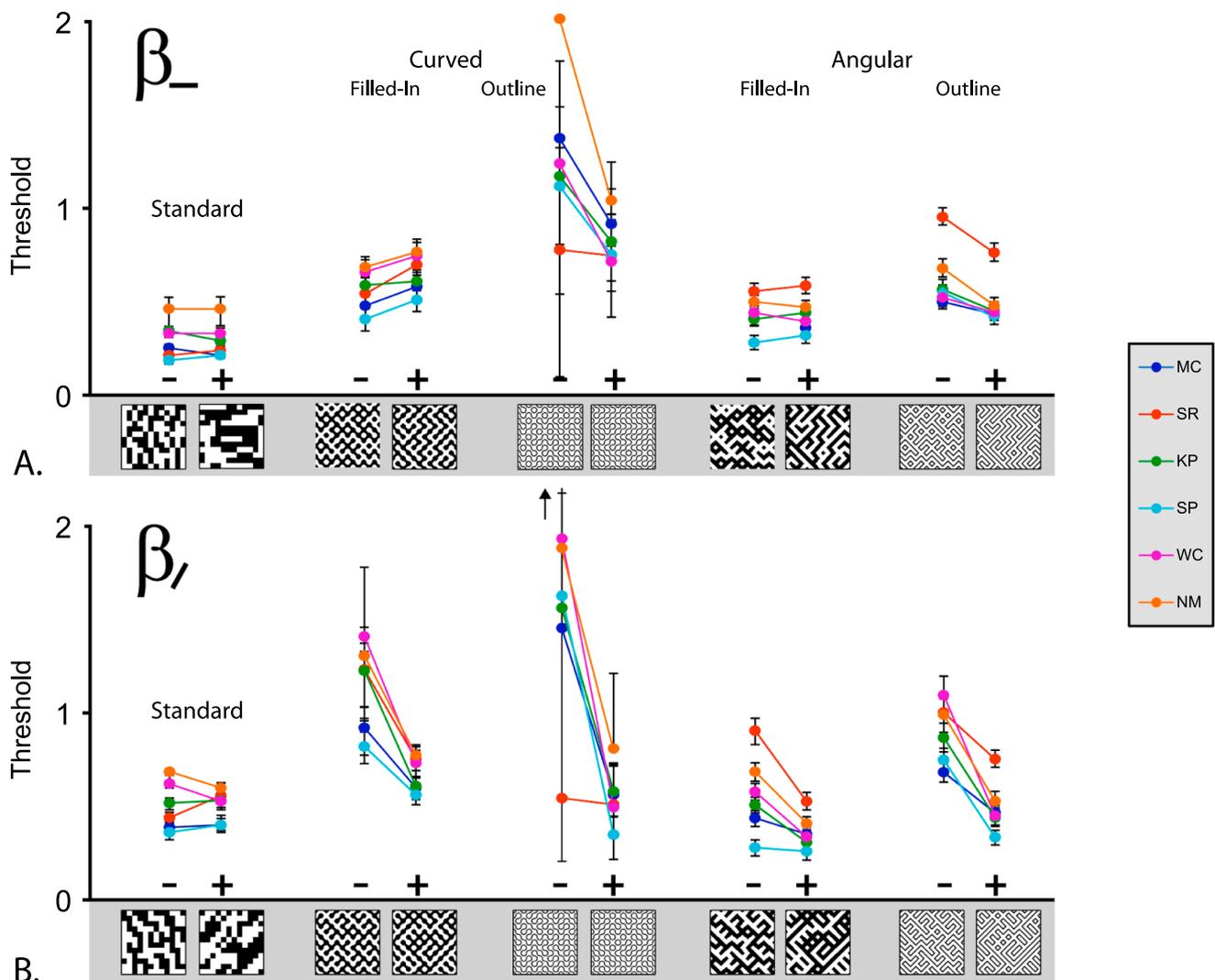


Fig. 6. Threshold data for second-order correlations ( $\beta_-$  and  $\beta_+$ ) for standard checks and the four token types. The data for standard checks and curved tokens are reproduced from Fig. 5. In one case ( $\beta_+$ , curved outline, negative correlation, subject WC), the upper end of the error bar is approximately 2.5, and is indicated by the upward arrow). Other plotting conventions as in Fig. 5. 6 subjects.

polarity (positive vs. negative); FvO: filled-in vs. outline tokens; and CvA: curved vs. angular tokens. All of these factors had statistically significant effects ( $p < 0.001$  after false-discovery-rate (FDR) correction). Two points demonstrate consistency across subjects: (a) the stimulus-related main factors had a much larger effect size (mean squares 1.72 to 3.47) than intersubject variation (mean square 0.22), and (b) the effect size for each of the four stimulus-related factors was more than 20 times the two-way interaction between these factors and Sbj.

The largest stimulus-related interaction is that of BeO  $\times$  Pol (mean square 0.76), and this is also highly consistent across subjects (Sbj  $\times$  BeO  $\times$  Pol, mean square 0.01); this corresponds to the observation that, across all token types, segmentation tends to be facilitated by  $\beta_- < 0$  and  $\beta_+ > 0$ . While this is the largest interaction of stimulus-related factors, all other two-, three-, and four-way interactions of stimulus factors are significant, with the exception of BeO  $\times$  FvO, and their effect sizes are all larger than the corresponding interactions with subject.

The four-way interaction (BeO  $\times$  Pol  $\times$  FvO  $\times$  CvA), which is consistent across subjects (across subject: within subject effect size ratio, 22) was highlighted above as indicating that the presence of circles is not the sole basis of the observed polarity-dependence. The nature of this interaction can be seen by building it up from its components, and

inspection of Fig. 6. Filled-in tokens yield BeO  $\times$  Pol interactions (i.e.,  $\beta_-$  yields higher threshold for positive than for negative polarity,  $\beta_+$  yields lower threshold for positive than for negative polarity; see second column of Fig. 6). But the interaction is changed for outline tokens (i.e.,  $\beta_-$  and  $\beta_+$  both yield lower thresholds for positive than for negative polarity; see third column of Fig. 6); this is the BeO  $\times$  Pol  $\times$  FvO interaction. This three-way interaction (BeO  $\times$  Pol  $\times$  FvO) is in turn modulated by FvA: curved tokens yield a larger three-way interaction than angular ones (fourth and fifth columns of Fig. 6).

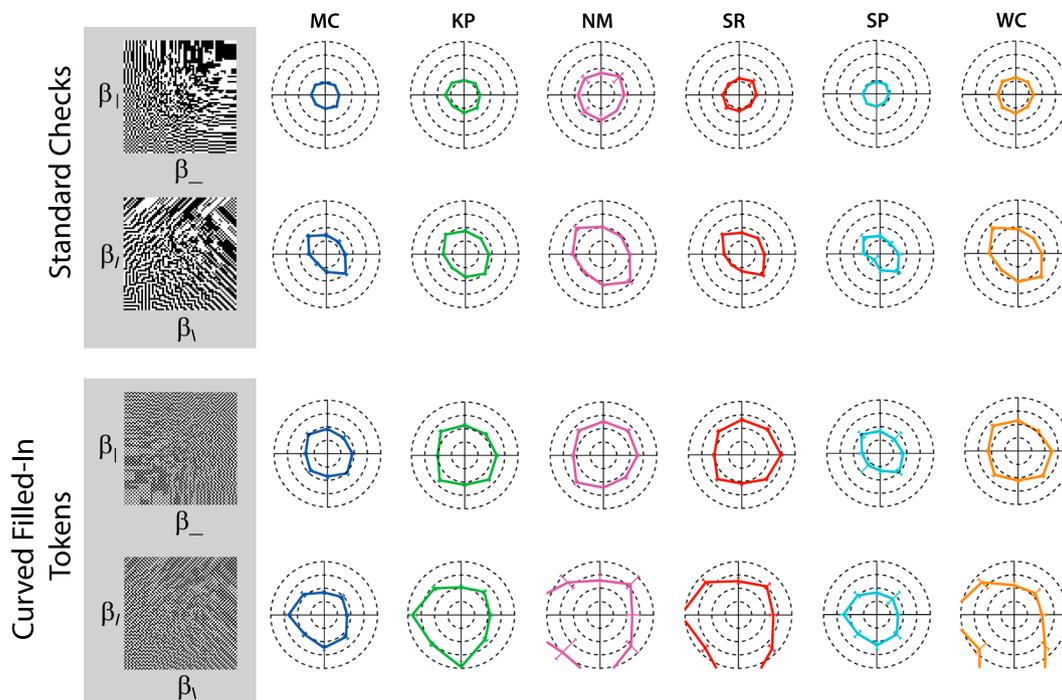
### 3.3. Experiment 3

Here we return to the original set of curved tokens, and examine how local second-order correlations interact. To do this, we constructed targets defined by combinations of either the two cardinal second-order correlations ( $\beta_-$ ,  $\beta_+$ ) or the two diagonal second-order correlations ( $\beta_1$ ,  $\beta_2$ ), and measured segmentation thresholds in several directions in these two coordinate planes (as in Victor and Conte (2012)).

Results are shown in Fig. 7. For comparison, thresholds for standard textures are shown at the top of the figure. As previously reported, the isodiscrimination contours are approximately elliptical and symmetric with respect to the origin, indicating that positive and negative

**Table 1**  
 Analysis of variance of the data of Experiment 2 (Fig. 6). Sources of variance are grouped according to stimulus-related factors (left, one degree of freedom each) or factors that involve the subject (middle, five degrees of freedom each). Effect size is quantified by the mean squared variation attributed to each source. Significance is quantified by an F-ratio; the raw p-value is given numerically and the false-discovery rate correction by asterisks. The two right-most columns compare the effect size across subjects (left) with the corresponding interaction with subject (middle). False-discovery-rate corrections are carried out separately for this column.

Sources of variance not involving Sbj				Sources of variance involving Sbj				ratio	
MS[1]	F(1,5)	p	MS[5]	F(5,5)	p	F(1,5)	p		
BeO	1.719	< 0.0001****	Sbj	0.223	< 0.0001****	209.6	< 0.0001****		
Pol	1.957	< 0.0001****	Sbj × BeO	0.008	0.0528	57.4	0.0006***		
FvO	2.699	< 0.0001****	Sbj × BeO × FvO	0.034	0.0024***	53.3	0.0008***		
CvA	3.466	< 0.0001****	Sbj × FvO	0.051	0.0009***	21.3	0.0057*		
BeO × Pol	0.756	< 0.0001****	Sbj × CvA	0.162	< 0.0001****	80.4	0.0003***		
BeO × FvO	0.000	0.6595	Sbj × BeO × Pol	0.009	0.0400*	0.0	0.8820		
BeO × CvA	0.026	0.0106*	Sbj × BeO × FvO	0.015	0.0155*	5.8	0.0606		
Pol × FvO	0.387	< 0.0001****	Sbj × BeO × CvA	0.005	0.1493	12.9	0.0157*		
Pol × CvA	0.086	0.0008***	Sbj × Pol × FvO	0.030	0.0032**	6.1	0.0572		
FvO × CvA	0.120	0.0004***	Sbj × Pol × CvA	0.014	0.0171*	2.5	0.1762		
BeO × Pol × FvO	0.049	0.0029**	Sbj × FvO × CvA	0.048	0.0011***	3.2	0.1342		
BeO × Pol × CvA	0.063	0.0016***	Sbj × BeO × Pol × FvO	0.015	0.0146*	18.1	0.0081*		
BeO × FvO × CvA	0.075	0.0011***	Sbj × BeO × CvA	0.004	0.2143	57.7	0.0006***		
Pol × FvO × CvA	0.037	0.0054**	Sbj × BeO × FvO × CvA	0.001	0.6177	4.6	0.0858		
BeO × Pol × FvO × CvA	0.036	0.0055**	Sbj × Pol × FvO × CvA	0.008	0.0542	21.8	0.0055*		
abbreviations mean square [degrees of freedom] Sbj subject (6) BeO orientation of second-order correlation (2) Pol polarity (positive vs. negative) of correlation (2) FvO filled-in vs. outline tokens (2) CvA curved vs. angular tokens (2)									
FDR-corrected significance * p < 0.05 ** p < 0.01 *** p < 0.005 **** p < 0.001									



**Fig. 7.** Isodiscrimination contours for textures constructed with combinations of two cardinal second-order correlations ( $\beta_-$ ,  $\beta_+$ ) or two diagonal second-order correlations ( $\beta_1$ ,  $\beta_2$ ). Top: textures are constructed with standard black and white checks. Bottom: textures are constructed with curved filled-in tokens. Error bars indicate 95% confidence limits by bootstrap. 6 subjects. The image on the left of each row shows the texture domain defined by variation of the correlation strengths over the range  $-1$  to  $+1$ .

correlations are equally strong cues for segmentation, and that sub-threshold statistical cues combine in an approximately quadratic fashion (thresholds for MC, KP, and SR previously reported in (Victor, Rizvi, et al., 2017); thresholds for NM, SP, and WC are new measurements, carried out in parallel with the current measurements for curved tokens).

However, this simple behavior does not hold when the textures are constructed with curved tokens, rather than black and white checks. The asymmetries along the coordinate axes corroborate the findings of Experiment 1 (Fig. 5, middle panels) and Experiment 2 (Fig. 6, curved filled-in tokens): thresholds for negative values of the cardinal  $\beta_-$ 's ( $\beta_-$  and  $\beta_+$ ) are lower than for positive values for all subjects, while thresholds for negative values of the diagonal  $\beta_-$ 's ( $\beta_1$  and  $\beta_2$ ) are higher than for positive values for all subjects.

In the ( $\beta_-$ ,  $\beta_+$ )-plane, there also are asymmetries along the  $\beta_- = \beta_+$ -diagonal, and these asymmetries were typically opposite to the asymmetries along the coordinate axes (i.e., thresholds for  $\beta_-$  and  $\beta_+$  simultaneously negative are higher than when both are positive, in 4 of 6 subjects). The difference between the asymmetries along the axes (mean positive:negative ratio, 1.16) and along the main diagonal (mean positive:negative ratio, 0.96) was significant ( $p < 0.001$ , two-tailed paired  $t$ -test). In the ( $\beta_1$ ,  $\beta_2$ )-plane, the most pronounced asymmetries are along the axes, and the thresholds in either the positive or negative direction along  $\beta_1 = \beta_2$  are close to equal. The net effect of these asymmetries is that for textures constructed with tokens, the isodiscrimination contours deviate substantially from ellipses, indicating that the underlying cues combine in a more complex fashion than for standard textures.

#### 4. Discussion

This study analyzes aspects of form that enable the segmentation of a visual image into component textures, and, like many previous studies (Bergen & Julesz, 1983; Caelli & Julesz, 1978; Caelli, Julesz, & Gilbert, 1978; Chubb et al., 2004; Chubb, Nam, Bindman, & Sperling, 2007; Graham, 1989; Graham & Sutter, 1998; Julesz, 1981; Landy & Oruc,

2002; Nothdurft & Li, 1985; Saarela & Landy, 2012; Victor, Rizvi, et al., 2017; Victor et al., 2015), see Victor, Conte, and Chubb (2017) for a recent review), views the texture segmentation paradigm as a way to assay the kinds of local features that the visual system extracts. However, in contrast to much previous work (Chubb et al., 2004, 2007; Graham & Sutter, 1998; Landy & Oruc, 2002; Nothdurft & Li, 1985; Saarela & Landy, 2012; Victor, Rizvi, et al., 2017; Victor et al., 2015), here the focus is not on elementary features such as luminance, contrast, and orientation, but on features that are more shape-like – the presence of small closed figures and convexity. To ensure that we are measuring the impact of these latter features, stimuli are equated for luminance, contrast, and orientation content.

To generate such stimuli, we build on the texture space introduced in Victor and Conte (2012), a 10-dimensional domain whose coordinates specify the correlations among  $2 \times 2$  neighborhoods of black or white checks. This domain allowed for the manipulation of the occurrence of edges and corners, but not the aspects of shape considered here. By replacing the uniform checks used in those stimuli by tokens containing appropriately-designed contours, the original coordinates of the space now modulate the formation of elongated contours and closed shapes, without changing the number of more elementary features. Replacing the uniform checks by tokens leads to a substantial change in the sensitivity to the underlying correlations (e.g., Fig. 5), indicating that performance is not driven by these underlying correlations *per se*, but on the visual features that they create.

##### 4.1. Local image statistics and the features they induce

A central aspect of our approach is the stimulus set: it is designed to allow for manipulation of the occurrence of elements of form (elongated contours and simple shapes), while keeping many lower-level features constant. Specifically, variation of the second- and higher-order correlations used to create the underlying texture does not change the probability of either kind of check, and therefore, does not change the distribution of local oriented contours, or the average contrast or

luminance. Moreover, since the tokens are designed so that there are no discontinuities where they meet, there are no terminators, T-junctions, or X-junctions (Bergen & Julesz, 1983) in any of the token stimuli.

As a first step in interpreting our results, we determine the relationship between the values of the second- and higher-order correlations, and the probability with which specific configurations are formed. We begin by considering the configurations that have the potential to be seen as convex boundaries, as this involves only two checks. Such configurations occur when tokens that share an edge have opposite orientations – forming semicircles in the case of curved tokens, or half of a diamond in the case of angular tokens (Fig. 2 row 2). The probability that horizontally-adjacent tokens contain opposite orientations is given by  $\frac{1}{2}(1 - \beta_-)$ ; for vertically-adjacent tokens, it is  $\frac{1}{2}(1 - \beta_v)$ . Thus, the probability that a pair of adjacent checks (either horizontally-adjacent or vertically-adjacent) forms a contour that may be seen as convex is given by

$$p_{convex} = \frac{1}{2} \left( \frac{1}{2}(1 - \beta_-) + \frac{1}{2}(1 - \beta_v) \right) = \frac{1}{4}(2 - \beta_- - \beta_v) \quad (2)$$

Note that if two adjacent checks have the same orientation (Fig. 2 row 1), then they form either a gentle S-shaped curve (for curved tokens) or a straight line (for angular tokens), rather than a potential convexity.

We next consider the probability that the four checks that form a  $2 \times 2$  neighborhood create a closed shape (a circle for curved tokens or a diamond for the angular tokens) – as shown in Fig. 2 row 5. To analyze the situation, we use “0” to indicate a token whose diagonal runs from lower left to upper right, and “1” to indicate a token whose diagonal runs from upper left to lower right. With this convention, a closed shape is formed by four nearest-neighbor checks in a  $2 \times 2$  neighborhood if they are in the configuration  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . The probability of this configuration depends not only on the  $\beta_-$ 's, but, because four checks are involved, on  $\alpha$  as well. It is given by Table 1 of Victor and Conte (2012):

$$p_{closed} = p \left( \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) = \frac{1}{16}(1 - 2\beta_- - 2\beta_v + \beta_+ + \beta_\alpha + \alpha) \quad (3)$$

In both Eqs. (2) and (3), the probability of occurrence of a local feature is a linear function of the texture coordinates. This is a consequence of the way that the texture coordinates are set up (Victor & Conte, 2012): the probability of any configuration within a  $1 \times 2$ ,  $2 \times 1$ , or  $2 \times 2$  region is a linear function of the texture coordinates.

The linear dependence of feature occurrence on the texture coordinates has implications for interpreting the shapes of the isodiscrimination contours found in Experiment 3. In our previous studies as well as studies in other visual perceptual domains (Macadam, 1942; Nothdurft, 2000; Poirson & Wandell, 1990; Saarela & Landy, 2012), subthreshold signals from different features combine in an approximately quadratic fashion. If these features are in turn linear functions of the coordinates, then isodiscrimination contours will be quadratic functions of the coordinates and will therefore have elliptical shapes. This prediction holds for the standard textures (Fig. 7 top), but not for textures composed of curved tokens (Fig. 7 bottom), especially in the  $(\beta_v, \beta_-)$ -plane. Such a breakdown implies either that subthreshold signals combine in a manner that is not approximately quadratic, or, that there are important visual features that are defined by regions larger than  $2 \times 2$  (since the occurrence of such features will not be linear functions of the coordinates). We suspect that the latter is the more likely, given the qualitative success of quadratic-combination laws in many domains, as noted above.

#### 4.2. Small closed contours

Recognizing that the existence of longer-range features will necessarily qualify our conclusions, we now use Eqs. (2) and (3) to interpret the results of Experiments 1 and 2. We first consider the

diagonal second-order correlations  $\beta_\lambda$  and  $\beta_j$ . Of all the features considered above (elongated contours, convex elements, small closed contours, and peninsulas), they only modulate the occurrence of small closed contours, with positive values of  $\beta_\lambda$  or  $\beta_j$  increasing their occurrence (Eq. (3)). Positive values of these coordinates lead to reduced thresholds for segmentation, independent of whether tokens are curved or angular, and independent of whether they are filled-in or outlined (Figs. 5 and 6). The most parsimonious explanation for these findings is that segmentation is facilitated in textures with a greater occurrence of small closed contours.

A potential confound is that positive values of  $\beta_\lambda$  or  $\beta_j$  also increase the probability that two diagonally-related checks have the same orientations (Fig. 2 row 2). Arguing against this explanation is the observation that positive values of the fourth-order correlation  $\alpha$  – which increase the occurrence of closed contours but do not change the probability that diagonally-related checks have the same orientation – also lowers the segmentation threshold (Fig. 5, bottom). Thus, independent of whether nearby checks have coherent orientations or not, as well as the other low-level features mentioned above, our data indicate that the presence of small closed contours lead to lower segmentation thresholds. Also, as mentioned above, modulation of  $\beta_\lambda$  or  $\beta_j$  induce changes in the configurations of neighborhoods larger than  $2 \times 2$ , but it seems unlikely that these features would play a role while smaller features would not.

Asymmetries in the role of features (including “closure”) are well-known to be prominent in visual search (Hulleman et al., 2000; Treisman & Gormican, 1988; Treisman & Souther, 1985; Wolfe, 2001). Rosenholtz (2001) and Rosenholtz, Huang, Raj, Balas, and Ilie (2012) have argued that these asymmetries may not reflect asymmetries in the fundamental visual computations, but rather in the experimental design: stimuli consist of one target among many distractors. However, here we find asymmetry in a comparison of performance in a segmentation task in which either the target or background is enriched in a feature, vs. a segmentation task in which one or the other is depleted of that feature. Thus, the experimental design is the same for increases vs. decreases of a feature, and not likely to be the source of the asymmetry in performance.

#### 4.3. Convexity

The effect of the cardinal second-order correlations  $\beta_-$  and  $\beta_v$  is more complex. In contrast to the effect of the diagonal second-order correlations, which is independent of token type (Fig. 6B), the effect of cardinal second-order correlations depends on the token (Fig. 6A). For curved filled-in tokens, positive values increase threshold, but when the curved regions are replaced by outlines, positive values result in a decreased threshold. Note that Eqs. (2) and (3) show that positive values of  $\beta_-$  and  $\beta_v$  have two effects: they decrease the occurrence of segments that may be seen as convex (replacing them by segments whose orientations are coherent), and they decrease the occurrence of closed contours.

We interpret this dependence on whether the tokens are filled-in vs. outlined as indicating that convexity, rather than just curvature, is the relevant feature. For a curved contour, convexity vs. concavity can only be determined after first resolving whether the object that owns the contour is on one side or the other. Textures composed of outline tokens do not appear to be composed of objects, so there can be little sense of border ownership, and hence, of convexity or concavity. When the tokens are filled-in, either the white or the black regions can be perceived as the object. However, there is a preference for the percept in which the object is convex (Bertamini & Wagemans, 2013; Burge, Fowlkes, & Banks, 2010; Zhou, Friedman, & von der Heydt, 2000). The interpretation of the token-dependent effect of  $\beta_-$  and  $\beta_v$  as implicating convexity is also supported by the observation that it is stronger for curved tokens than for angular ones.

While the primary effects observed (the token-independent effects

of the diagonal second-order correlations, and the token-dependent effects of the cardinal second-order correlations) can be accounted for parsimoniously in terms of closed contours and convexity, our results cannot be taken as evidence against a role for other features, such as concavity (Barenholtz et al., 2003; Cohen, Barenholtz, Singh, & Feldman, 2005; Hulleman et al., 2000). Wherever a convexity is formed by two tokens of opposite orientations that share an edge, a concavity can be produced if the border-ownership is reversed. More generally, distinguishing between convexity and concavity is difficult if not impossible, as it requires appealing to border ownership (Bertamini & Wagemans, 2013), and, in many paradigms, a range of controls to eliminate potential confounds due to changes in lower-level aspects of the stimulus (Cohen et al., 2005; Hulleman et al., 2000). Inflection points could also play a role, as they are formed at an edge whenever two curved tokens share opposite orientations, i.e., wherever the local contour is S-shaped, rather than either concave or convex. However, since inflection points are independent of border ownership, this feature would not lead to a difference between filled-in and outline tokens.

#### 4.4. Filter-rectify-filter models

Finally, we note that while our approach, which is based on maximum-entropy textures, is rather different than the envelope-and-carrier stimuli used by many others (e.g., Hallum et al. (2011), Landy and Oruc (2002), Sutter et al. (1995) and Westrick et al. (2013), there are clear points of contact. Such textures are ideally suited for testing filter-rectify-filter (FRF) models, as the carrier probes the characteristics of the first filter, while the envelope probes the properties of the second. While FRF models are unlikely to account for extraction of features based on multipoint correlations (Victor & Conte, 1991) or features such as closed contours, they clearly can extract statistical structure based on pairwise correlations. This holds both for the underlying standard textures and derived textures based on tokens. In the latter case, it is worth noting that the textures constructed by Westrick et al. (2013) (their Fig. 2B and C) have many of the characteristics of the curved filled-in token textures with  $\beta_- > 0$ , and this suggests how a FRF model could account for some of the findings here.

#### Acknowledgements

Portions of this work were presented at the 2016 meeting of the Vision Sciences Society (St. Petersburg, FL). This work was supported by NIH NEI EY7977.

#### References

Adelson, E. H. (2001). On seeing stuff: the perception of materials by humans and machines. *Proceedings of the SPIE, human vision and electronic imaging VI*, (pp. 1–12).

Barenholtz, E., Cohen, E. H., Feldman, J., & Singh, M. (2003). Detection of change in shape: An advantage for concavities. *Cognition*, 89(1), 1–9.

Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature*, 303, 696–698.

Bertamini, M., & Wagemans, J. (2013). Processing convexity and concavity along a 2-D contour: Figure-ground, structural shape, and attention. *Psychonomic Bulletin & Review*, 20(2), 191–207. <https://doi.org/10.3758/s13423-012-0347-2>.

Burge, J., Fowlkes, C. C., & Banks, M. S. (2010). Natural-scene statistics predict how the figure-ground cue of convexity affects human depth perception. *Journal of Neuroscience*, 30(21), 7269–7280. <https://doi.org/10.1523/JNEUROSCI.5551-09.2010>.

Caelli, T., & Julesz, B. (1978). On perceptual analyzers underlying visual texture discrimination: Part I. *Biological Cybernetics*, 28(3), 167–175.

Caelli, T., Julesz, B., & Gilbert, E. (1978). On perceptual analyzers underlying visual texture discrimination: Part II. *Biological Cybernetics*, 29(4), 201–214.

Chubb, C., Landy, M. S., & Economou, J. (2004). A visual mechanism tuned to black. *Vision Research*, 44(27), 3223–3232.

Chubb, C., Nam, J. H., Bindman, D. R., & Sperling, G. (2007). The three dimensions of human visual sensitivity to first-order contrast statistics. *Vision Research*, 47(17), 2237–2248. <https://doi.org/10.1016/j.visres.2007.03.025>.

Cohen, E. H., Barenholtz, E., Singh, M., & Feldman, J. (2005). What change detection tells us about the visual representation of shape. *Journal of Vision*, 5(4), 313–321. <https://doi.org/10.1167/5.4.3>.

Freeman, J., Ziemba, C. M., Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (2013). A functional and perceptual signature of the second visual area in primates. *Nature Neuroscience*, 16(7), 974–981. <https://doi.org/10.1038/nn.3402>.

Graham, N. (1989). *Visual pattern analyzers*. Oxford: Clarendon Press.

Graham, N., & Sutter, A. (1998). Spatial summation in simple (Fourier) and complex (non-Fourier) texture channels. *Vision Research*, 38(2), 231–257.

Hallum, L. E., Landy, M. S., & Heeger, D. J. (2011). Human primary visual cortex (V1) is selective for second-order spatial frequency. *Journal of Neurophysiology*, 105(5), 2121–2131. <https://doi.org/10.1152/jn.01007.2010>.

Hermundstad, A. M., Briguglio, J. J., Conte, M. M., Victor, J. D., Balasubramanian, V., & Tkacik, G. (2014). Variance predicts salience in central sensory processing. *Elife*, 3. <https://doi.org/10.7554/eLife.03722>.

Hulleman, J., te Winkel, W., & Boselie, F. (2000). Concavities as basic features in visual search: Evidence from search asymmetries. *Percept Psychophys*, 62(1), 162–174.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802), 91–97.

Julesz, B., Gilbert, E. N., & Victor, J. D. (1978). Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics*, 31(3), 137–140.

Karklin, Y., & Lewicki, M. S. (2009). Emergence of complex cell properties by learning to generalize in natural scenes. *Nature*, 457(7225), 83–86. <https://doi.org/10.1038/nature07481>.

Kingdom, F. A., Keeble, D., & Moulden, B. (1995). Sensitivity to orientation modulation in micropattern-based textures. *Vision Research*, 35(1), 79–91.

Landy, M. S., & Oruc, I. (2002). Properties of second-order spatial frequency channels. *Vision Research*, 42(19), 2311–2329.

Macadam, D. L. (1942). Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America*, 32, 247–273.

Nirenberg, S., & Pandarinath, C. (2012). Retinal prosthetic strategy with the capacity to restore normal vision. *Proceedings of the National Academy of Sciences of the United States of America*, 109(37), 15012–15017. <https://doi.org/10.1073/pnas.1207035109>.

Nothdurft, H. C. (2000). Salience from feature contrast: Additivity across dimensions. *Vision Research*, 40(10–12), 1183–1201.

Nothdurft, H. C., Gallant, J. L., & Van Essen, D. C. (2000). Response profiles to texture border patterns in area V1. *Visual Neuroscience*, 17(3), 421–436.

Nothdurft, H. C., & Li, C. Y. (1985). Texture discrimination: Representation of orientation and luminance differences in cells of the cat striate cortex. *Vision Research*, 25(1), 99–113.

Poirson, A. B., & Wandell, B. A. (1990). The ellipsoidal representation of spectral sensitivity. *Vision Research*, 30(4), 647–652.

Rosenholtz, R. (2001). Search asymmetries? What search asymmetries? *Percept Psychophys*, 63(3), 476–489.

Rosenholtz, R., Huang, J., Raj, A., Balas, B. J., & Ilie, L. (2012). A summary statistic representation in peripheral vision explains visual search. *Journal of Vision*, 12(4), <https://doi.org/10.1167/12.4.14>.

Rust, N. C., Schwartz, O., Movshon, J. A., & Simoncelli, E. P. (2005). Spatiotemporal elements of macaque v1 receptive fields. *Neuron*, 46(6), 945–956.

Saarela, T., & Landy, M. S. (2012). Combination of texture and color cues in visual segmentation. *Vision Research*, 58, 59–67.

Sharan, L., Liu, C., Rosenholtz, R., & Adelson, E. H. (2013). Recognizing Materials using Perceptually Inspired Features. *International Journal of Computer Vision*, 103(3), 348–371. <https://doi.org/10.1007/s11263-013-0609-0>.

Sutter, A., Sperling, G., & Chubb, C. (1995). Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision Research*, 35(7), 915–924.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8(2), 194–214.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95(1), 15–48.

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, 114(3), 285–310.

Victor, J. D., Chubb, C., & Conte, M. M. (2005). Interaction of luminance and higher-order statistics in texture discrimination. *Vision Research*, 45(3), 311–328.

Victor, J. D., & Conte, M. M. (1991). Spatial organization of nonlinear interactions in form perception. *Vision Research*, 31(9), 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, Optics, Image Science, and Vision*, 29(7), 1313–1345. <https://doi.org/10.1364/JOSAA.29.001313>.

Victor, J. D., Conte, M. M., & Chubb, C. F. (2017). Textures as probes of visual processing. *Annual Review of Vision Science*, 3, 275–296. <https://doi.org/10.1146/annurev-vision-102016-061316>.

Victor, J. D., Rizvi, S. M., & Conte, M. M. (2017). Two representations of a high-dimensional perceptual space. *Vision Research*. <https://doi.org/10.1016/j.visres.2017.05.003>.

Victor, J. D., Thengone, D. J., & Conte, M. M. (2013). Perception of second- and third-order orientation signals and their interactions. *Journal of Vision*, 13(4), 21. <https://doi.org/10.1167/13.4.21>.

Victor, J. D., Thengone, D. J., Rizvi, S. M., & Conte, M. M. (2015). A perceptual space of local image statistics. *Vision Research*, 117, 117–135. <https://doi.org/10.1016/j.visres.2015.05.018>.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin*, 138(6), 1172–1217. <https://doi.org/10.1037/a0029333>.

- Westrick, Z. M., Henry, C. A., & Landy, M. S. (2013). Inconsistent channel bandwidth estimates suggest winner-take-all nonlinearity in second-order vision. *Vision Research*, *81*, 58–68. <https://doi.org/10.1016/j.visres.2013.01.010>.
- Willmore, B. D., Prenger, R. J., & Gallant, J. L. (2010). Neural representation of natural images in visual area V2. *Journal of Neuroscience*, *30*(6), 2102–2114. <https://doi.org/10.1523/JNEUROSCI.4099-09.2010>.
- Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Percept Psychophys*, *63*(3), 381–389.
- Wolfson, S. S., & Landy, M. S. (1998). Examining edge- and region-based texture analysis mechanisms. *Vision Research*, *38*, 439–446.
- Yu, Y., Schmid, A. M., & Victor, J. D. (2015). Visual processing of informative multipoint correlations arises primarily in V2. *Elife*, *4*. <https://doi.org/10.7554/eLife.06604>.
- Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*(17), 6594–6611.