SELECTIVE VISUAL SENSITIVITY TO IMAGE STATISTICS IN A MULTIDIMENSIONAL TEXTURE SPACE

Jonathan D. Victor and Mary M. Conte

892.2

Department of Neurology and Neuroscience, Weill Cornell Medical College, New York, NY

Summary

Perception of real-world images depends on the processing of image statistics. They can drive segmentation of an image into objects, indicate the material composition of a surface, and provide the gist of a scene. To study how the visual system processes image statistics, we use visual textures, because this enables us to isolate individual statistics, and to measure their interactions.

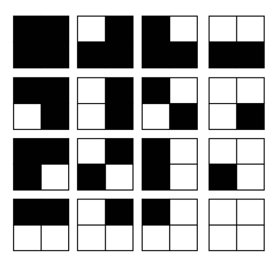
Image statistics constitute a high-dimensional set of parameters, including the luminance histogram (first order), the power spectrum (second order), and higher-order analogs. This leads to two related challenges: a large number of parameters must be specified to define a visual texture, and the set of image statistics cannot be explored exhaustively. To address the first challenge we define a texture by specifying only a small number of image statistics; the unspecified statistics are then determined implicitly, by creating textures that are as random as possible, given the specified parameters. To address the second challenge, we restrict consideration to a model "texture space" of black and white pixels, whose statistics are specified by the frequency of the colorings of 2x2 blocks of pixels. This class of statistics carries information about natural scenes (Tkačik et al., PNAS (2010)). As described below, this texture space is 10-dimensional, and its coordinates consist of first-, second-, third- and fourth-order image statistics, each of which is visually distinctive.

Using a simple segmentation task, we determined perceptual salience of the individual statistics: first-order statistics are more salient than second-order ones, but, interestingly, fourth-order statistics are more salient than third-order ones. For several pairs of statistics, we quantified their interactions. Isodiscrimination contours were approximately elliptical, indicating that image statistics are combined in an approximately quadratic fashion. These observations suggest two principles that may simplify the understanding of human perception of image statistics: selective processing of some kinds of statistics, and a Euclidean perceptual metric for texture space.

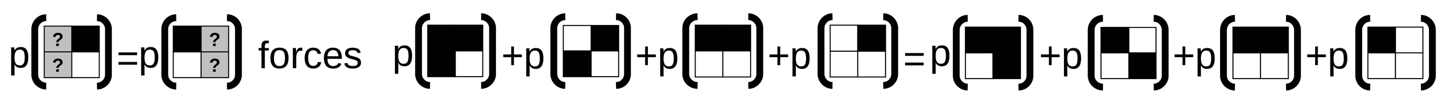
Coordinates for a Local Binary Texture Space

Why are there 10 parameters?

Textures are defined by the distribution of colorings of 2x2 blocks. There are 16 (=$2^{2\times2}$) possible colorings (right). However, there are constraints among these probabilities because the probabilities of smaller blocks cannot depend on *where* they occur within the 2x2 block.



For example,

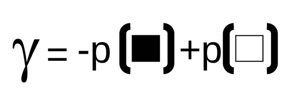


There are 6 independent linear constraints on 16 probabilities, yielding 16-6=10 free parameters.

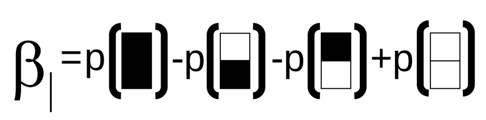
What are the 10 coordinates?

The 10 independent parameters are captured by 10 coordinates. Each is a linear combination of the probabilities. Each coordinate ranges from -1 to 1, and has a value of 0 for the random texture.

One coordinate for luminance:

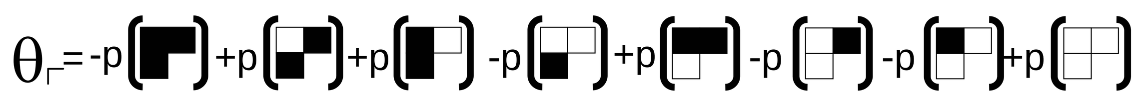


Four coordinates for pairwise correlation:



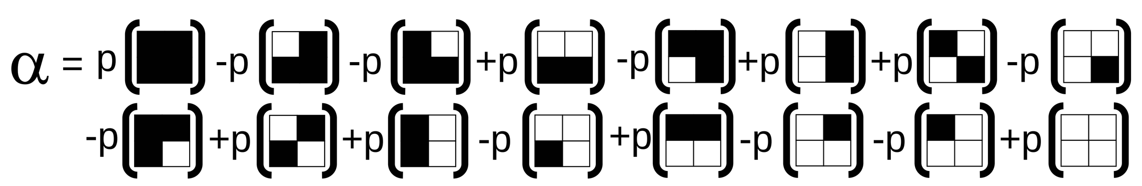
(similarly for β_{-}, $\beta_{|}$, $\beta_{\backslash}$, $\beta_{/}$)

Four coordinates for the parity of triplets:



(similarly for $\theta_{\lrcorner}$, $\theta_{\llcorner}$, $\theta_{\ulcorner}$, $\theta_{\urcorner}$)

One coordinate for the parity of quadruplets:

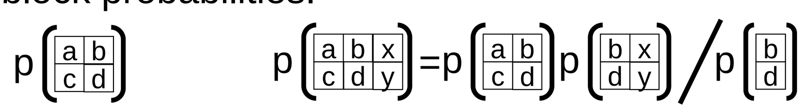


Texture Generation

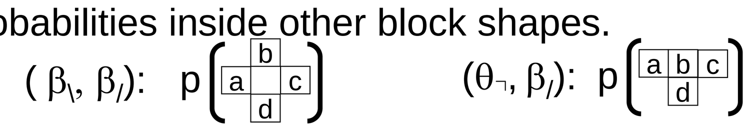
Basic algorithm: 2-D Markov processes

Step1. Assign 2x2 block probabilities.

Step 2. Propagate via the Markov rule.

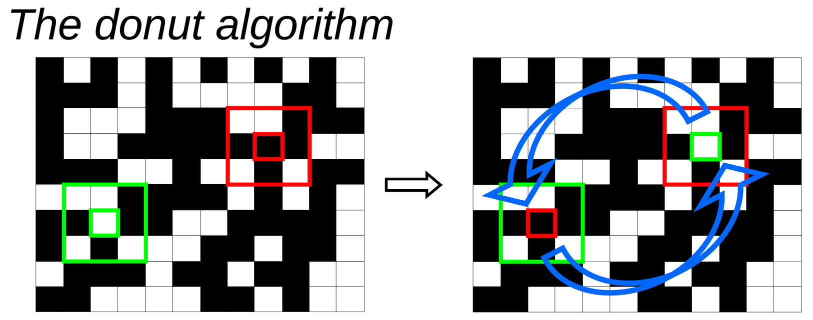


This yields a maximum-entropy texture for most coordinate pairs. For others, different algorithms are required. In one case ($\theta_{\lrcorner}$, $\theta_{\urcorner}$), we start with the standard 2x2 block; in other cases, we specify probabilities inside other block shapes.



In these cases, Markov propagation does not achieve maximum entropy. We therefore add a second step: iteratively swapping the interiors of matching "donuts." This preserves long-range correlations but does not change the 2x2 block probabilities, and achieves maximum entropy.

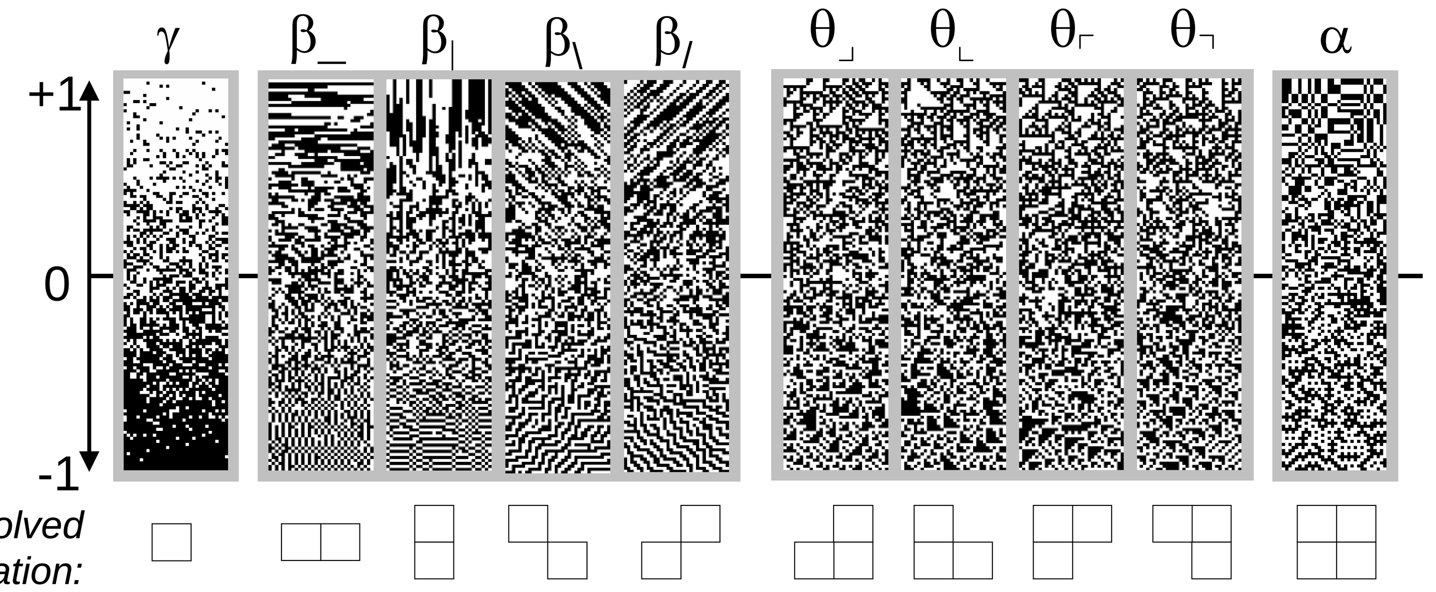
The donut algorithm



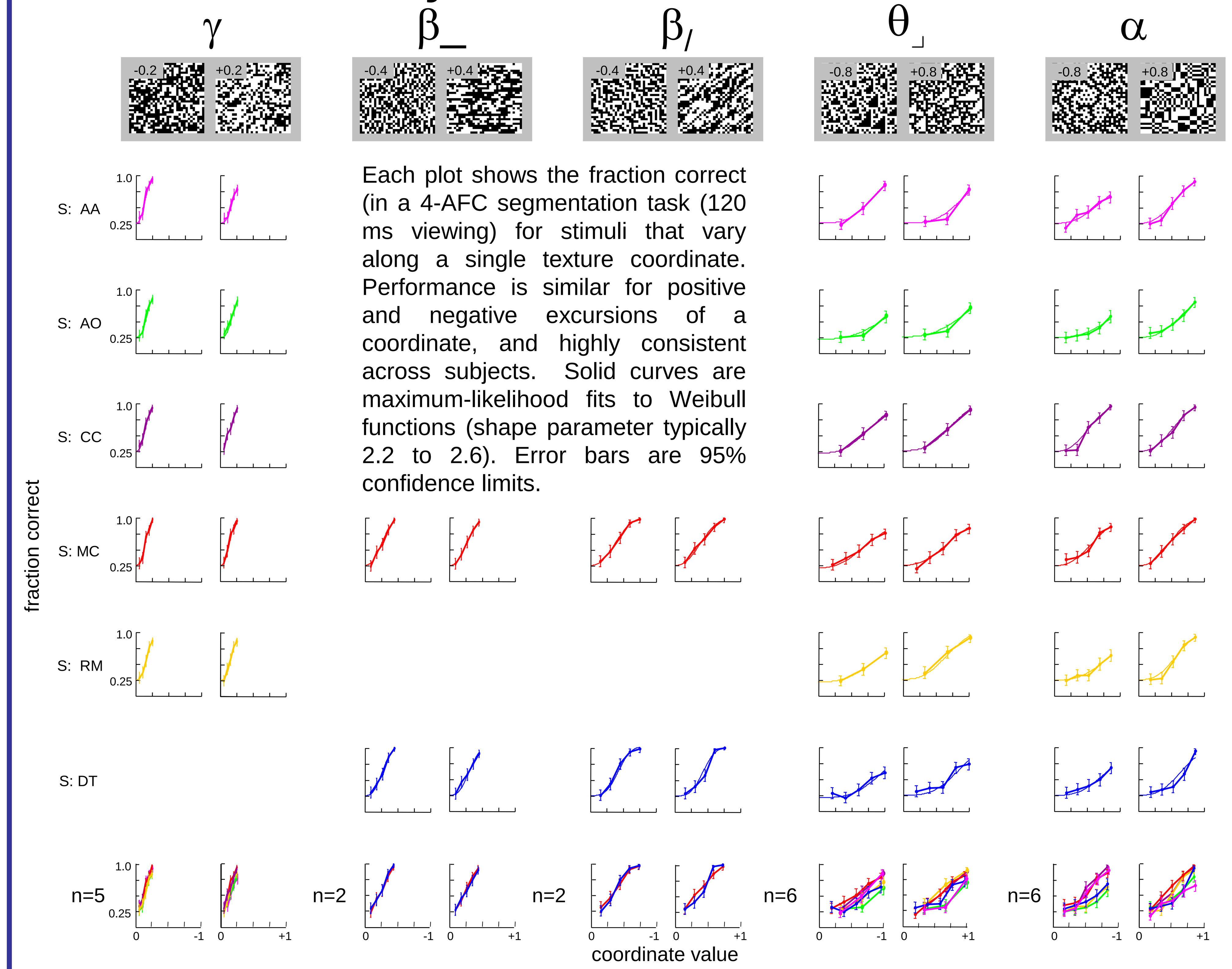
Coordinate Axes
selective sensitivity to individual statistics

Coordinate axes

Each strip shows the textures generated by varying one coordinate across its entire range, from -1 to 1. A coordinate value of 0 corresponds to a random texture.

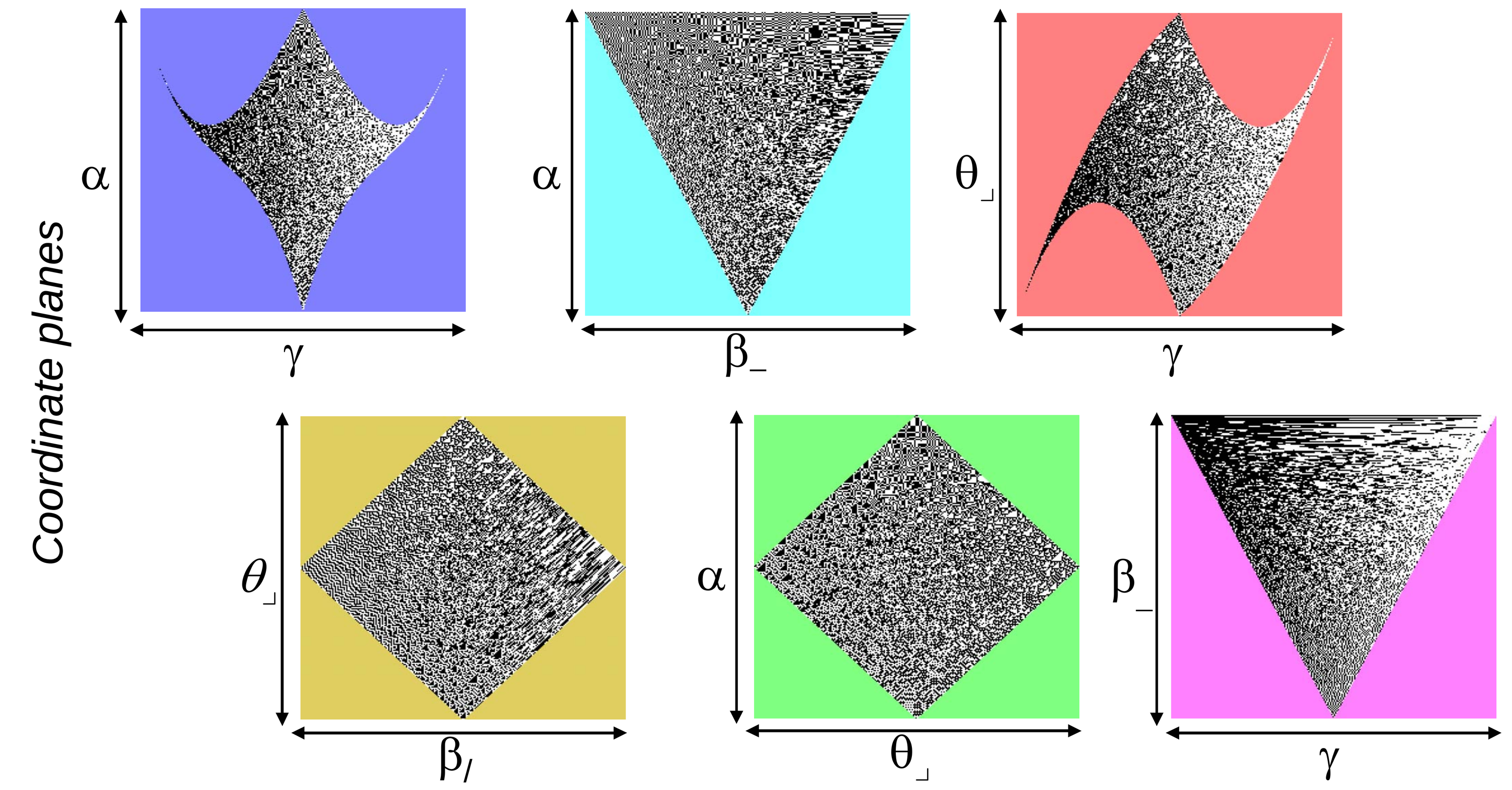


Psychometric Functions



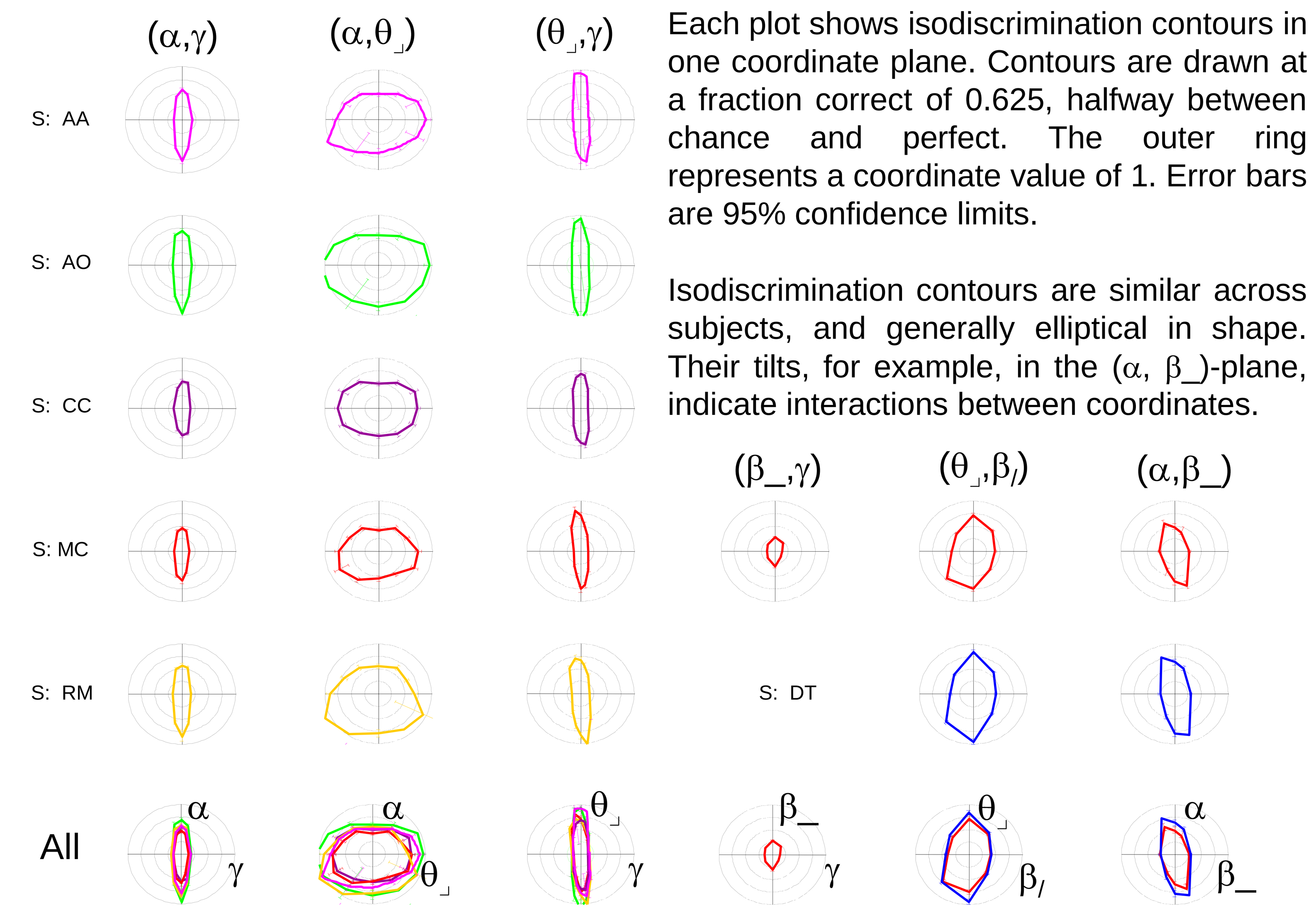
Each plot shows the fraction correct (in a 4-AFC segmentation task (120 ms viewing) for stimuli that vary along a single texture coordinate. Performance is similar for positive and negative excursions of a coordinate, and highly consistent across subjects. Solid curves are maximum-likelihood fits to Weibull functions (shape parameter typically 2.2 to 2.6). Error bars are 95% confidence limits.

Coordinate Planes
simple interactions between statistics



These images illustrate the texture coordinate pairs studied. The center of each image is random; the border indicates the limits imposed by consistency conditions. One can get an idea of the salience of each coordinate and how they interact by noting where each image becomes visibly non-random. This is quantified by isodiscrimination contours, shown below.

Isodiscrimination Contours



Each plot shows isodiscrimination contours in one coordinate plane. Contours are drawn at a fraction correct of 0.625, halfway between chance and perfect. The outer ring represents a coordinate value of 1. Error bars are 95% confidence limits.

Isodiscrimination contours are similar across subjects, and generally elliptical in shape. Their tilts, for example, in the (α, β_)-plane, indicate interactions between coordinates.

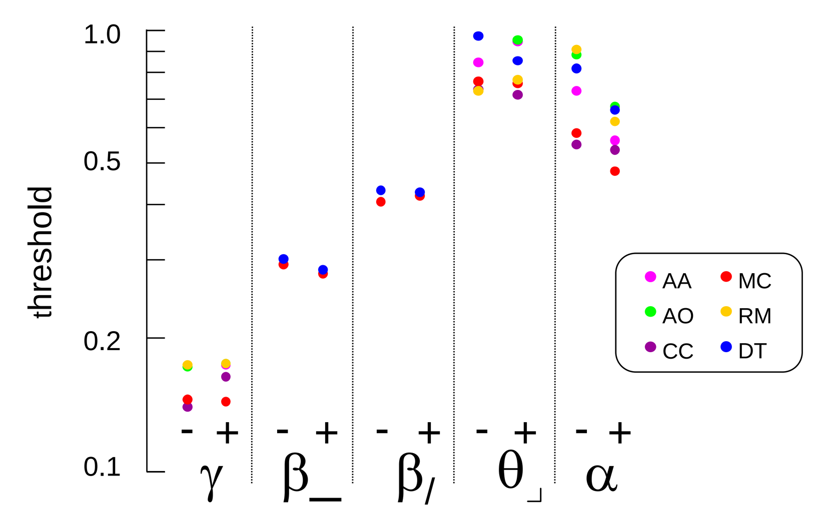
Psychophysical Methods

TASK
- Identify the location of the target stripe (4-AFC, top, right, bottom, left)
- Target is 16 x 64 pixels on a 64 x 64 pixel array
- Trials either have a structured target on a random background, or a random target on structured background

STIMULI
- Pixel Size: 14 min;
- Display Size: 14.8 deg [2],
- Binocular viewing at 1m
- Contrast 1.0
- Duration 120 ms (followed by mask)

SUBJECTS
- 6 subjects
- VA: 20/20, with correction if needed
- Practice: approx 1600 trials

CONDITIONS
- In the (γ,θ) - and (θ,α) -planes:
  - 8 repeats of 12 on-axis points
  - 8 repeats of 24 off-axis points
- In other planes:
  - 8 repeats of 20 on-axis points
  - 16 repeats of 8 off-axis points
- 288 trials per block, random order
- 15 blocks = 4320 trials per plane
- Feedback during practice only



Thresholds are highly consistent across subjects and are similar for positive and negative excursions, in approximate ratio 1:2:5:4 (γ:β:θ:α).

Conclusions

- Binary textures with local correlations form a perceptual space with 10 dimensions. Within this space, sensitivities to image statistics are highly consistent within and across observers.
- Pairwise interactions between image statistics are described by approximately elliptical isodiscrimination contours.
- Thresholds for first-, second-, third-, and fourth- order statistics are in the ratio of 1:2:5:4. We speculate that these relative sensitivities correspond to the extent to which the statistics are informative about natural images.

Supported by EY7977.

We thank Daniel Thengone for assistance with data collection and analysis.