# Characterization of visual sensitivity to image statistics in a multidimentional texture space


Mary M. Conte, Daniel J. Thengone, Jonathan D. Victor

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



799.06
SFN 2011

Supported by EY7977


## Motivation and Overview

The visual system is sensitive to many kinds of image statistics, including those of low and high order. In natural scenes, these different kinds of image statistics are intertwined in a complex fashion. This motivated the development of a library of synthetic images that enables testing the salience of individual image statistics and how they interact.

To reduce dimensionality, we consider only local statistics of binarized images. Based on the findings of Tkacik et al. (2010), we focus on the statistics of 2x2 neighborhoods. Correlations of all orders within these neighborhoods are completely described by 10 coordinates.

Here we fully characterize the sensitivity to each of these 10 image statistics and their pairwise interactions. Sensitivity is highly conserved across observers, and a Euclidean model of the perceptual space accurately accounts for the interactions. These findings define the geometry of a perceptual space based on the informative statistics of natural images.

## Methods

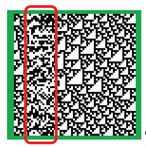
**SUBJECTS**
12 subjects
VA: 20/20, with correction if needed
Practice: approx 1600 trials

**CONDITIONS**
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

**STIMULI**
Pixel Size: 14 min
Display Size: 14.8 deg$^2$
Binocular viewing at 1m
Contrast: 1.0
Duration: 120 ms (followed by mask)
Target: 16 x 64 pixels on a 64 x 64 array
Trials either have a structured target on a random background, or random target on structured background

**TASK**
Find the location of the target stripe (4-AFC, top, right, bottom, left)

In this sample stimulus, the target stripe (defined by $\theta_{\lrcorner}$) is on the bottom on a random background.

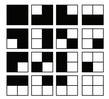


Here, the target stripe (on the left) is random on a structured ($\theta_{\lrcorner}$) background.

*Thresholds were not significantly different for these two conditions and did not depend on target location.*

## Coordinates for a Local Binary Texture Space

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.



Since there are 6 independent linear constraints on 16 probabilities, there are 16-6=10 free parameters., each of which corresponds to a coordinate.

One coordinate for luminance: $\gamma = -p(\blacksquare) + p(\square)$

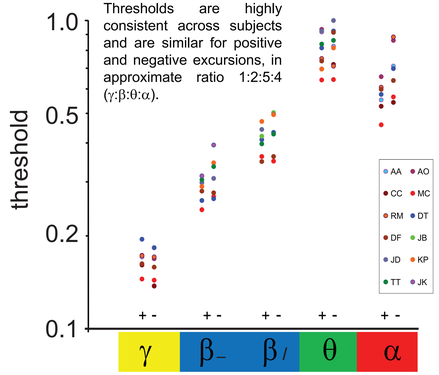
Four coordinates for pairwise correlation: $\beta_{|} = p(\blacksquare) - p(\blacksquare) - p(\blacksquare) + p(\square)$ (similarly for $\beta_{_}$, $\beta_{\backslash}$, $\beta_{/}$)

Four coordinates for the parity of triplets: $\theta_{\ulcorner} = -p(\blacksquare) + p(\blacksquare) + p(\blacksquare) - p(\blacksquare) + p(\blacksquare) - p(\blacksquare) - p(\blacksquare) + p(\square)$ (similarly for $\theta_{\lrcorner}$, $\theta_{\urcorner}$, $\theta_{\llcorner}$)

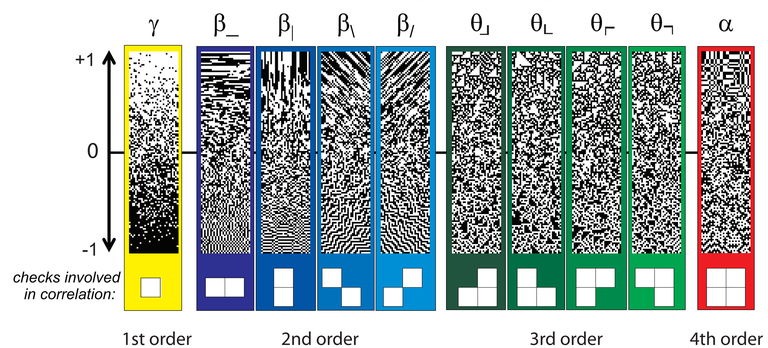
**One coordinate for the parity of quadruplets:** $\alpha = p(\blacksquare) - p(\blacksquare) - p(\blacksquare) + p(\blacksquare) - p(\blacksquare) + p(\blacksquare) + p(\blacksquare) - p(\square) - p(\blacksquare) + p(\blacksquare) + p(\blacksquare) - p(\blacksquare) + p(\blacksquare) - p(\blacksquare) - p(\blacksquare) + p(\square)$

### Sensitivity on Each Coordinate Axis

Thresholds are highly consistent across subjects and are similar for positive and negative excursions, in approximate ratio 1:2:5:4 ($\gamma$:$\beta$:$\theta$:$\alpha$).



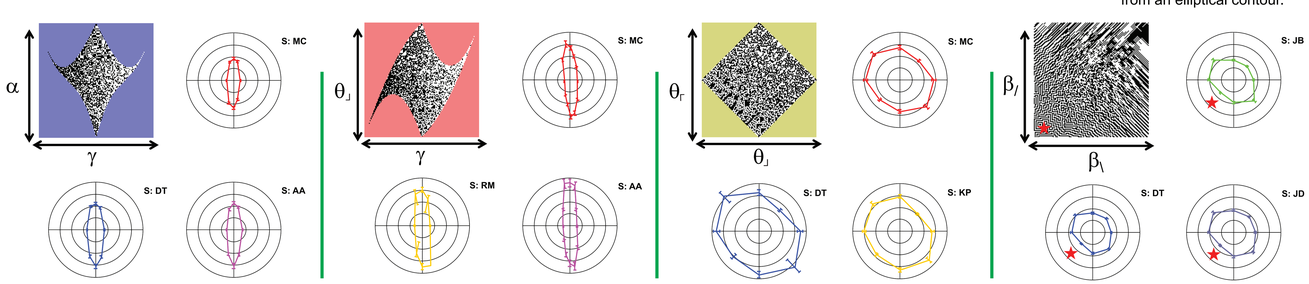

### Coordinate Axes




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.

## Isodiscrimination Contours in Selected Coordinate Planes

★ Note the consistent deviation from an elliptical contour.




## Thresholds and Model Fits in All 15 Unique Coordinate Planes

### Ellipsoid model for discrimination in the entire 10-d space

$$d^2 = \sum_{i,j} Q_{i,j} c_i c_j$$

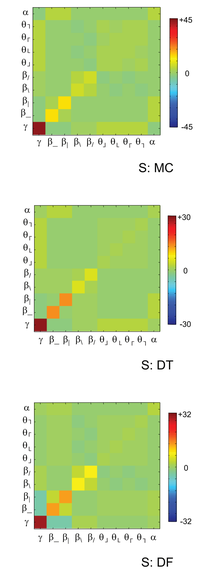
*d* is the perceptual distance from the random texture (*d*=1 corresponds to threshold)

$Q_{i,j}$ are the elements of a symmetric, positive-definite matrix

$c_i$ are the texture coordinates

Under the hypothesis of a Euclidean perceptual metric, the isodiscrimination surface is an ellipsoid in the 10-parameter texture space, described by a symmetric matrix *Q* (see equation). We constrained the elements $Q_{i,j}$ to be equal for image statistics that differed only by a 90 deg. rotation. The resulting model was highly over-determined: it has 5 unique diagonal parameters and 15 unique off-diagonal parameters, and there are 128 different threshold values to account for.

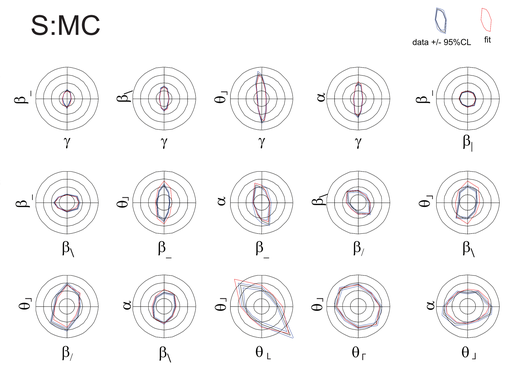
Fitted values of $Q_{i,j}$ were consistent across subjects (see color maps).

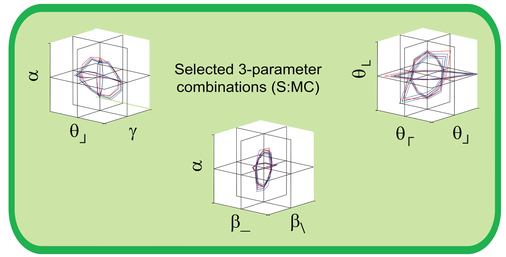


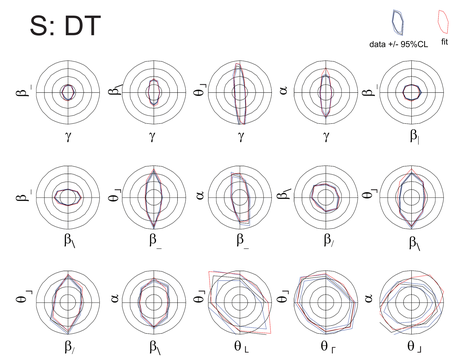

Goodness of Fit (rmse)

| Subject | Raw | Corrected |
|---|---|---|
| MC | 0.065 | 0.031 |
| DT | 0.086 | 0.031 |
| DF | 0.091 | 0.047 |
| Mean | 0.081 | 0.036 |

The raw rmse is the root-mean-squared deviation between the predictions of the ellipsoid model and the measured thresholds. The corrected rmse compares the predictions of the ellipsoid model with the best possible predictions of any opponent model.

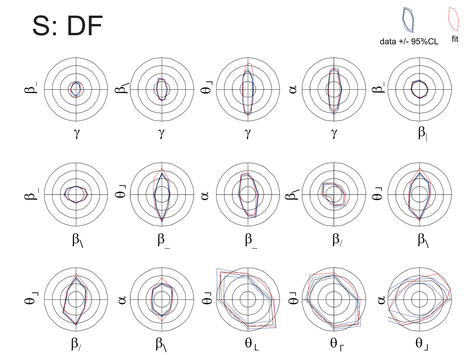



The planar plots show all of the experimentally-determined thresholds in each of the coordinate planes tested, along with 95% confidence limits (via bootstrap), and fits of the Euclidean model.

Some predicted contours deviate slightly from ellipses because they take into account the values of out-of-plane parameters.




## Summary and Conclusions

- Local image statistics of low and high order can be described by a 10-parameter stimulus space.
- The salience of individual statistics and their interactions is highly conserved across observers.
- A Euclidean perceptual metric accounts for > 90% of the variance.
- Deviations from the Euclidean model, though small, are largely consistent across observers. They include the greater sensitivity to positive vs. negative values for second- and fourth-order image statistics, and an interaction unique to two second-order statistics, $\beta_{\backslash}$ and $\beta_{/}$.


Reference: Tkačik, G., Prentice, J., Victor, J.D., and Balasubramanian, V. (2010) Local statistics in natural scenes predict the saliency of synthetic textures. *PNAS* USA 107, 18149-18154.