A model of sensitivity to binary local image statistics: Testing the predictions

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

Brain & Mind Research Institute, Weill Cornell Medical College, NY

mmconte@med.cornell.edu

Supported by EY7977

Motivation and Overview

Early cortical stages of visual analysis rely on processing of local correlations, as these define the elements (lines, edges, and texture) that must be extracted for scene segregation and object identification. In natural images, these correlations are of low and high order, and occur together in complex mixtures. To analyze how they are processed, we synthesize image sets in which they vary independently – thus generating a 10-dimensional "texture space". Stimuli drawn from this space enable characterization of perceptual sensitivities to many kinds of individual image statistics, alone and in combinations.

We made measurements of perceptual sensitivities along all 10 axes of the space, and in planes determined by their pairs. In each of the planes, the contours are nearly elliptical, and the shapes and orientations of the ellipses were similar across subjects. The elliptical contour shape suggests that sensitivity in the full 10-d space is described by an ellipsoid. Since the 10-d ellipsoid is uniquely determined by in-plane measurements, the model predicts visual sensitivity to complex combinations of image statistics. As we show, these predictions are accurate.

Methods and Psychometric Functions

SUBJECTS
6 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

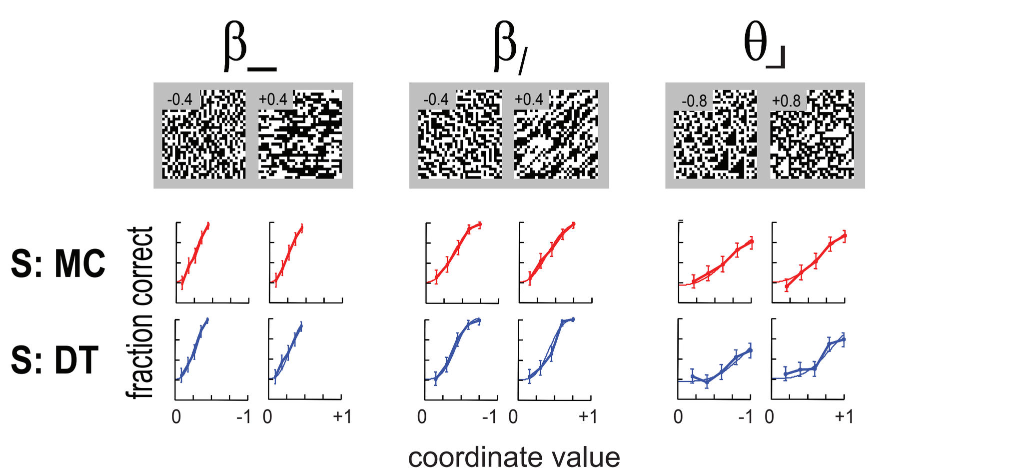
STIMULI
Pixel Size: 14 min
Display Size: 14.8 deg^2
Binocular viewing at 1m
Contrast: 1.0
Duration: 120 ms (followed by mask)
Target: 16 x 64 pixels on a 64 x 64 array
Trials either have a structured target on a random background, or random target on structured background

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.

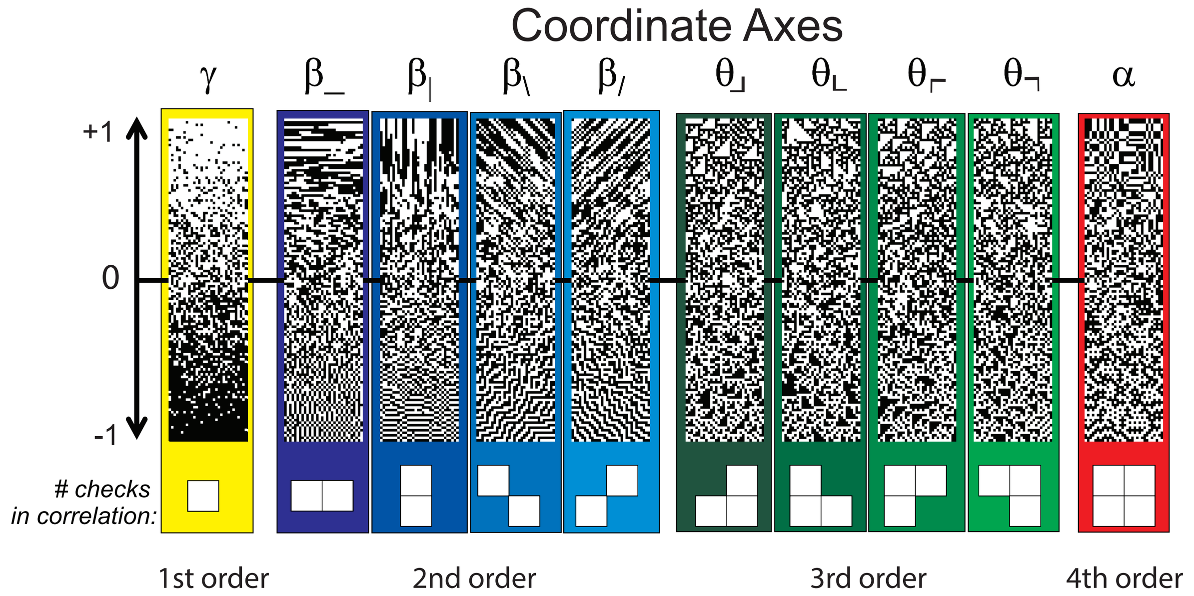
Here, the target stripe (on the left) is random on a structured ($\theta_{\lrcorner}$) background.

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



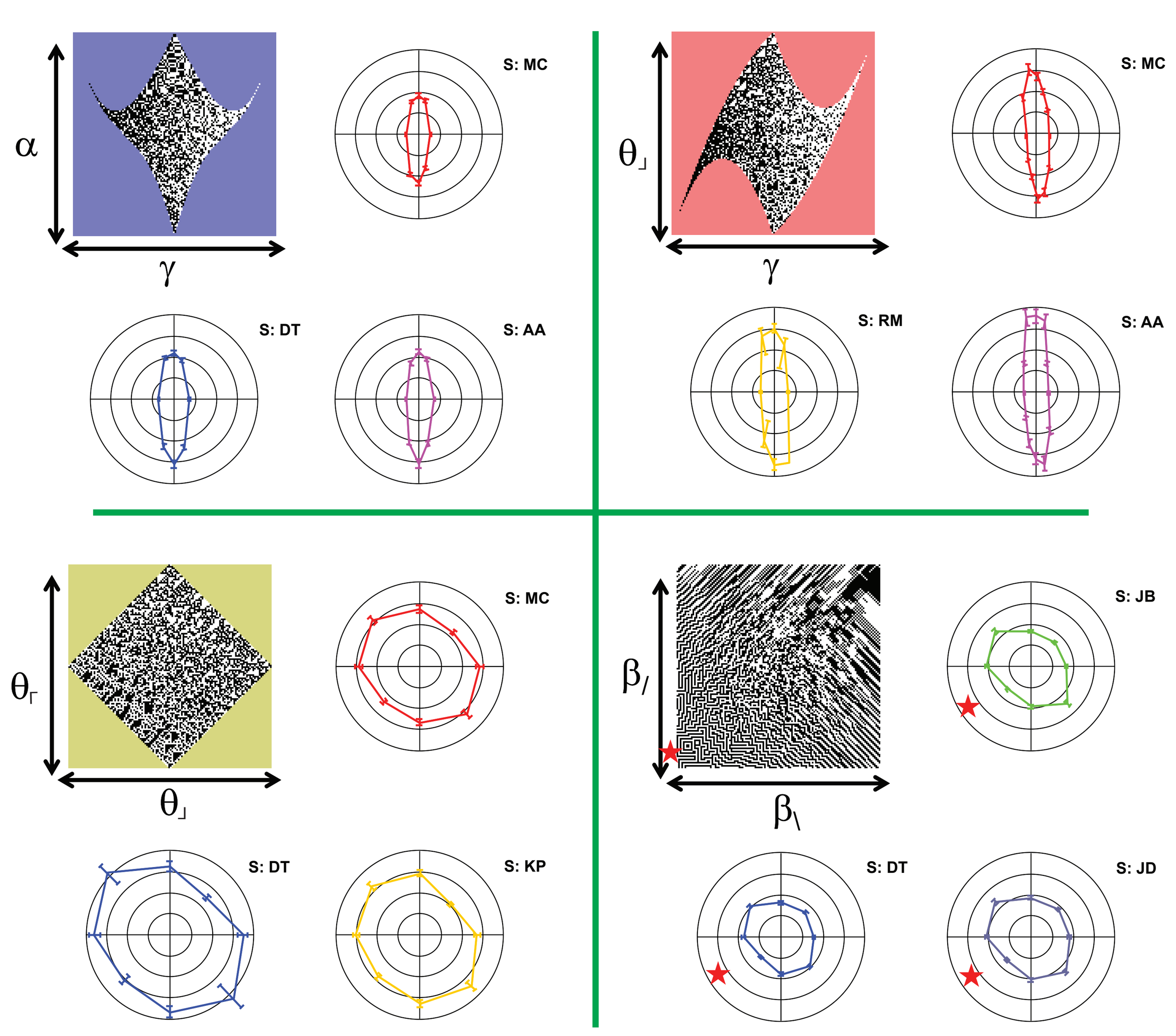
Plots show fraction correct for stimuli that vary along a single texture coordinate. Performance is similar for positive and negative excursions of a coordinate, and was highly consistent across subjects (MC and DT). Curves are maximum-likelihood fits to Weibull functions (shape parameter range: 2.2 - 2.6). Error bars are 95% confidence limits.

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.

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

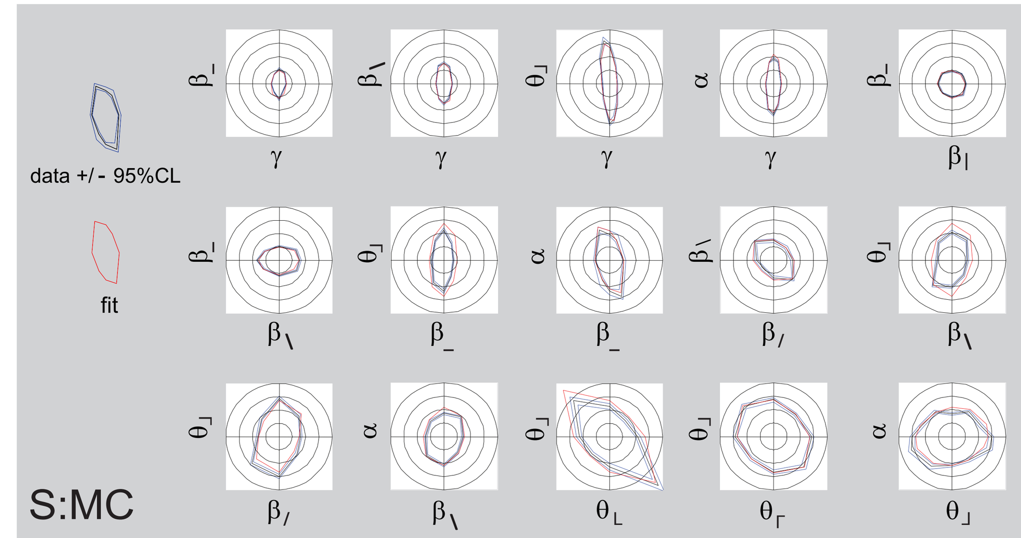
Goodness of Fit (rmse)

Subject	Raw	Corrected
MC	0.065	0.031
DT	0.086	0.031
DF	0.091	0.047
JD	0.099	0.043
Mean	0.085	0.038

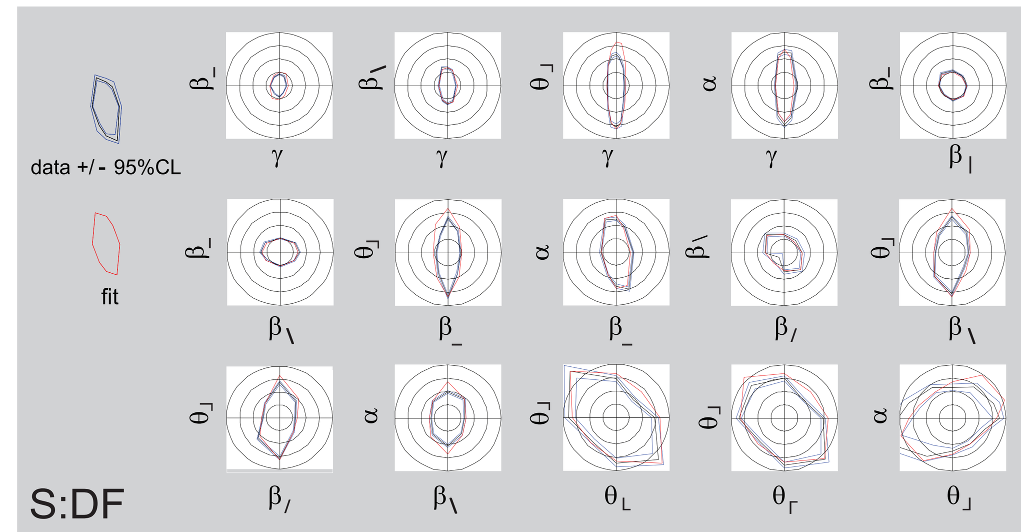
The raw rmse is the root-mean-squared deviation between the predictions of the ellipsoid model and the measured thresholds. The corrected rmse compares the predictions of the ellipsoid model with the best possible predictions of any opponent model, i.e., any model that predicts equal thresholds for positive and negative deviations of the image statistics.

The ellipsoid model and psychophysical data are in good agreement for each subject. We next use the ellipsoid model to compare across subjects (next column), and then show that it correctly predicts sensitivities to directions in the texture space not used to fit the model (final column).

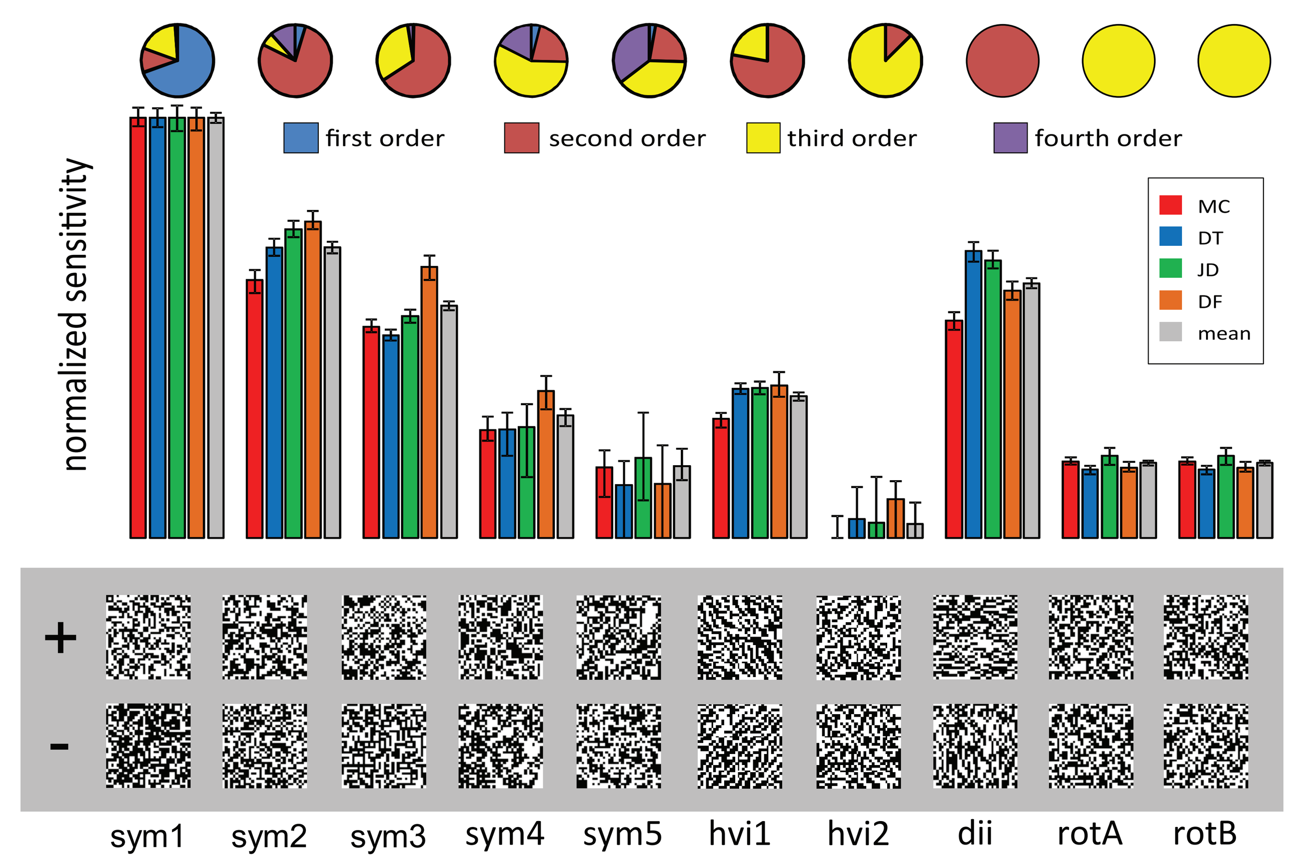
Thresholds and Model Fits in All 15 Unique Coordinate Planes



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

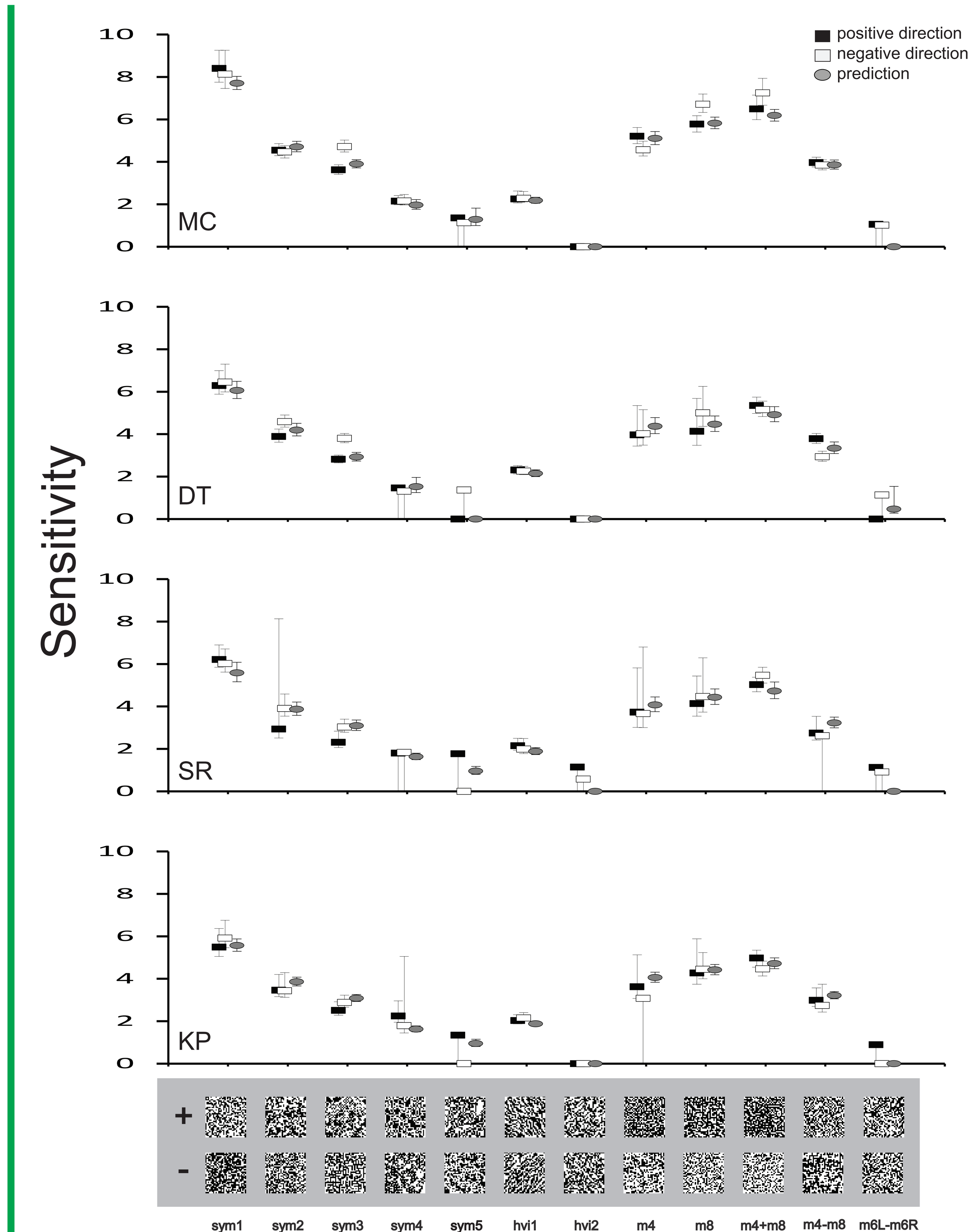


Consistency Across Subjects



Example textures representing the 10 principal axes. On each axis, texture samples are shown in both directions from the origin (which corresponds to the random texture). Correlations strengths, which correspond to distance from the origin, are 0.18 (for sym1 and sym2), and 0.36 (for the remaining textures). The pie charts show the contributions of first-, second-, third-, and fourth-order correlations to each direction.

Ellipsoid Model Predicts Sensitivities in Multiple Directions in the Texture Space



For each subject, model prediction and measured sensitivities are shown for 12 directions. The first seven directions are principal axes of the ellipsoid. (Only seven principal axes are tested, because the other three axes are in the coordinate planes, and therefore not out-of-sample.) The last five directions are Minkowski directions, which correspond to textures that have maximum or minimum porosity, as quantified by the number of holes per unit area.

CONCLUSIONS

- Psychophysical sensitivity to image statistics of low and high order can be modeled by ellipsoidal isodiscrimination contours.
- This model accurately predicts sensitivities to combinations of image statistics, including combinations that are predicted to be maximally salient, and combinations predicted to be undetectable.