Local image statistics: A highly conserved perceptual space encompassing statistics of low and high order

Jonathan D. Victor, Daniel T. Thengone, Charles F. Chubb, Mary M. Conte

Brain and Mind Research Institute, Weill Cornell Medical College

56.535
VSS 2013

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 (Victor & Conte, 2012).

To simplify the problem, we consider only local statistics in binary images. Based on the findings of Tkačik et al. (2010), we focus on the statistics of 2x2 neighborhoods. This yields 10 distinct image statistics that range in order from 1 to 4 and capture image features including luminance, contrast, orientation, blobs, edges, and corners.

Here, we fully characterize the sensitivity to each of these 10 image statistics and their pairwise combinations. The psychophysical data are well-described by an ellipsoidal isodiscrimination surface. We then use this characterization to show that sensitivities are highly conserved across subjects, both in terms of the orientation and magnitude of this ellipsoid's axes.

Methods and Psychometric Functions

SUBJECTS
13 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

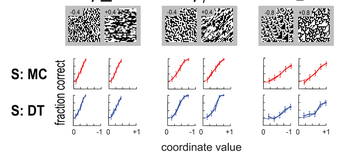
STIMULI
Pixel Size: 14 min
Display Size: 14.8 deg^2
Binocular viewing at 1m
Contrast: 1.0
Duration: 120 ms (followed by mask)
Target: 16 x 64 pixels on a 64 x 64 array
Trials either have a structured target on a random background, or random target on structured background

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

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

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

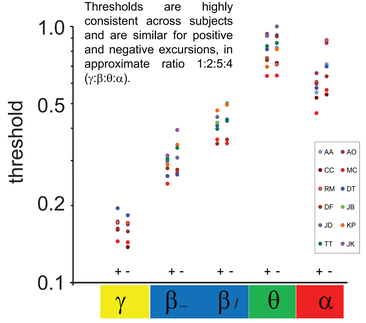
Thresholds were not significantly different for these two conditions and did not depend on target location.



Each plot shows the 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 (shown: MC and DT). Curves are maximum-likelihood fits to Weibull functions (shape parameter typically 2.2 to 2.6). Error bars are 95% confidence limits.

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 (γ:β:θ:α).



Coordinates for a Local Binary Texture Space

Textures are defined by the distribution of colorings of 2x2 blocks. There are 16 (=2^{2x2}) 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.

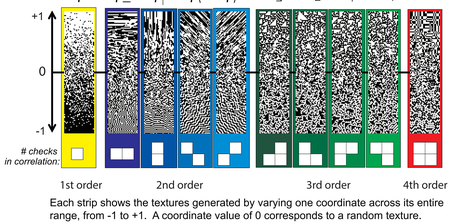
One coordinate for luminance: $\gamma = -p(\square) + p(\blacksquare)$

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

Four coordinates for the parity of triplets: $\theta_{\square} = -p(\square) + p(\square) + p(\square) - p(\square) - p(\square) + p(\square) + p(\square) - p(\square)$ (similarly for $\theta_{\_}$, $\theta_{\backslash}$, $\theta_{/}$)

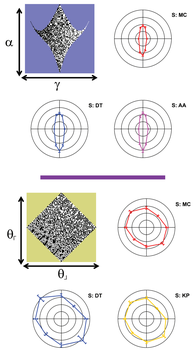
One coordinate for the parity of quadruplets: $\alpha = p(\square) - p(\square) - p(\square) + p(\square) + p(\square) - p(\square) - p(\square) + p(\square) - p(\square) + p(\square) + p(\square) - p(\square) + p(\square) - p(\square) - p(\square) + p(\square)$

Coordinate Axes

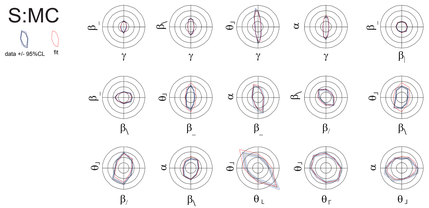


Each strip shows the textures generated by varying one coordinate across its entire range, from -1 to +1. A coordinate value of 0 corresponds to a random texture.

Isodiscrimination Contours in Selected Coordinate Planes



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

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

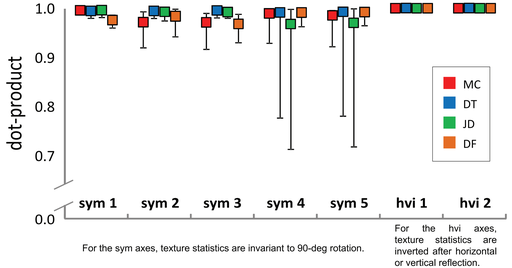
Goodness of Fit (rmse)

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.

Because the ellipsoid model and psychophysical data are in good agreement, we next use the ellipsoid model to compare across subjects: we first compare the directions of the axes of the ellipsoids, and then compare their magnitudes. We also show that the inferred sensitivities along the principal axes are in good agreement with measured thresholds.

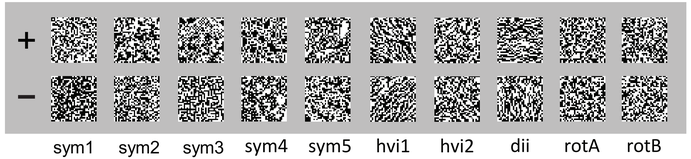
Consistency Across Subjects



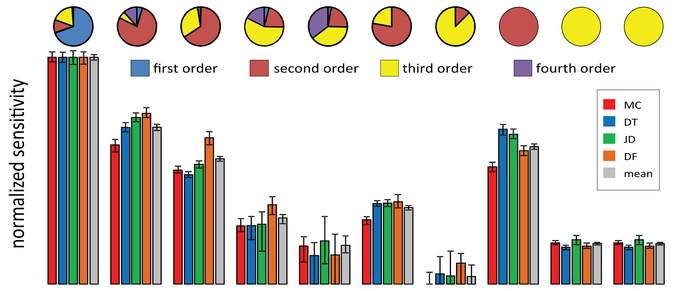
For the sym axes, texture statistics are invariant to 90-deg rotation.

For the hvi axes, texture statistics are inverted after horizontal or vertical reflection.

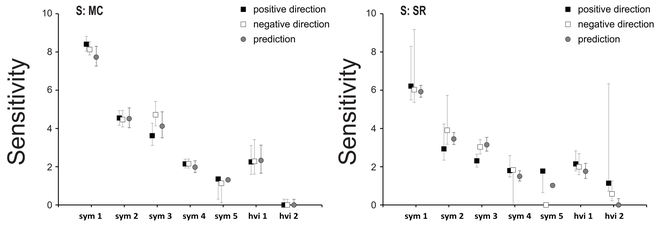
Directions of the principal axes are consistent across subjects. Directional consistency is quantified by the dot product between the directions determined for each subject (N=4), and the average direction. Note that although there are 10 principal axes (in the 10-dimensional texture space), only seven principal axes are compared; the directions of the other three axes are forced by symmetry to be identical 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).



Sensitivities along the 10 principal axes are consistent across subjects. Pies show how individual image statistics contribute to the 10 principal axes. Bars show sensitivity along these axes, for N=4 subjects and their mean. The axes shown here include the 7 axes above (sym1 to sym5 and hvi1 to hvi2) and three other axes for which the principal directions are determined by symmetry, but the sensitivities are experimentally measured for each subject. Note that for the axis hvi2, the ellipsoid model predicts that the sensitivity is zero, i.e., that in this direction in texture space, psychophysical performance is at chance.



Predicted and measured sensitivities for the 7 out-of-sample principal axes of the isodiscrimination surface.

Conclusion

Local binary image statistics constitute a 10-dimensional space whose perceptual metric is highly conserved across subjects. Predicted sensitivities along the principal axes of the isodiscrimination surface are confirmed by measurement.

Supported by EY7977