Towards a model for sensitivity to local image statistics

Jonathan D. Victor, Syed M. Rizvi, Jacob W. Bush, Mary M. Conte

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

VSS 2017
53.4105

Segmenting an image into its components is a crucial step in visual processing. Modeling the underlying computations is challenging, owing to the high dimensionality of the problem. We therefore have taken a reductionist approach, focusing on black-and-white images defined by their local spatial correlations. In this space, image statistics can be independently probed and fully analyzed, enabling a predictively accurate model of psychophysical sensitivities in 10 dimensions.

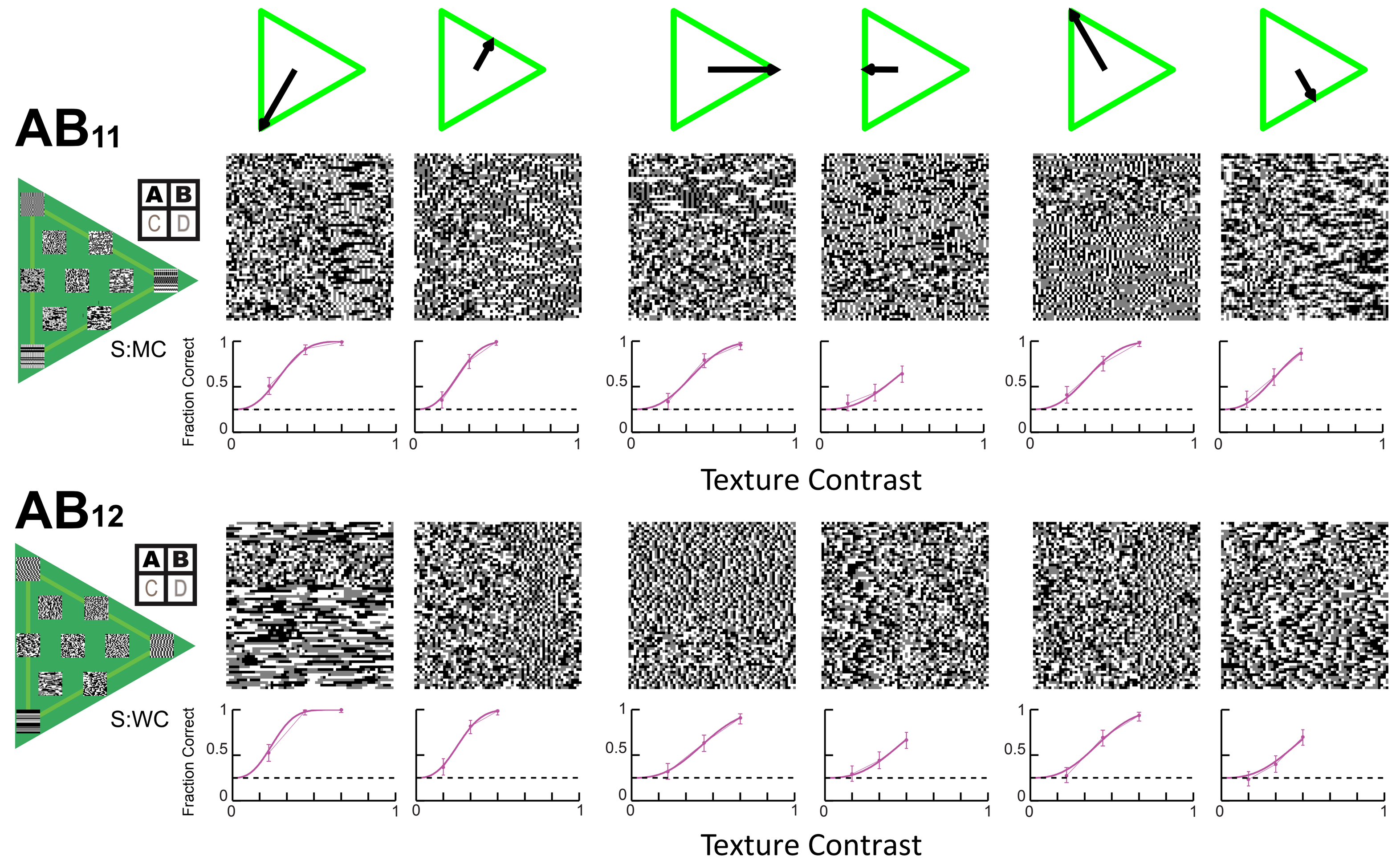
Here, we show that key aspects of these findings extend to images with three luminance levels – a space with 66 dimensions. We suggest a model framework that can account for the observed behavior. This framework is also consistent with Chubb et al.'s studies of "scramble" textures with multiple gray levels.

Stimulus Domains and Psychometric Functions

The stimulus domain is parameterized by the probabilities of all configurations of black, gray, and white checks in 2x2 patches. For ternary textures, there are 66 dimensions (2 first-order, 16 second-order, 32 third-order, and 16 fourth-order). We study a subset of these dimensions.

To parameterize the domain, the four checks in a 2x2 patch are designated A, B, C, and D. The parameters AB_{XY}, for example, describe how the luminance values L(A) and L(B) are correlated. Specifically, the parameter AB_{XY} indicates the distribution of values of q=X*L(A)+Y*L(B) (mod 3), where black = 0, gray = 1, and white = 2.

Each parameter spans a triangular gamut (right) that specifies the distribution of q. The vertices are the textures for which q is maximally biased, i.e., always 0, always 1, or always 2. The centroid is the random texture, where q is equally likely to be 0, 1, or 2.



Fraction correct in two second-order planes (AB_{11}, AB_{12}) for subjects MC, and WC. In each plane, 6 directions are shown (the directions towards and away from each vertex, as indicated by the arrows). The maximum on the abscissa is a texture contrast of 1: the distance from the random texture to a texture of maximum structure.

Stimulus examples are shown above each curve. Error bars are 95% confidence limits. Smooth curves are fits to Weibull functions, and the texture contrast yielding a fraction correct of 0.625 was taken as threshold. Note that performance is similar for positive variations of an image statistic (towards a vertex) and negative variations (away from a vertex).

Methods

Subjects
N = 6
VA: 20/20

Conditions
4-AFC Design
288 trials per block
15 blocks per condition
4,320 trials per plane
423,360 total trials

Stimuli
Contrast: 1.0
Check size: 14 min
Display size: 14.8 deg ^{2}
Binocular viewing at 1 m
Duration: 120 ms followed by a 300 ms mask
Target: 16 x 64 pixels on a 64 x 64 array
Practice: 50 trials per condition
Feedback during practice only

Support: EY7977

References

Chubb, C., Landy, M.S. and Econopouly, J. (2004). A visual mechanism tuned to black. Vision Research **44**, 3223-3232.

Chubb, C., Nam, J., Bindman, D.R., and Sperling, G. (2007). The three dimensions of human visual sensitivity to first-order statistics. Vision Research **47**, 2237-2248.

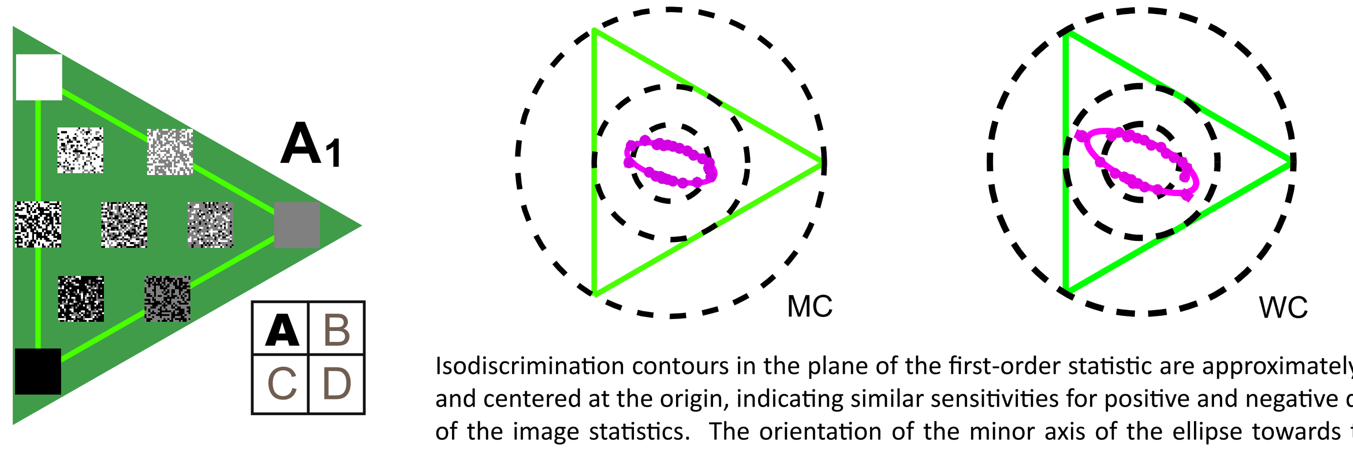
Silva, A.E. and Chubb, C. (2014). The 3-dimensional, 4-channel model of human visual sensitivity to grayscale scrambles. Vision Research **101**, 94-107.

Victor, J.D. and Conte, M.M. (2012) Local image statistics: maximum-entropy constructions and perceptual salience. J Opt Soc Am **29**, 1313-1345.

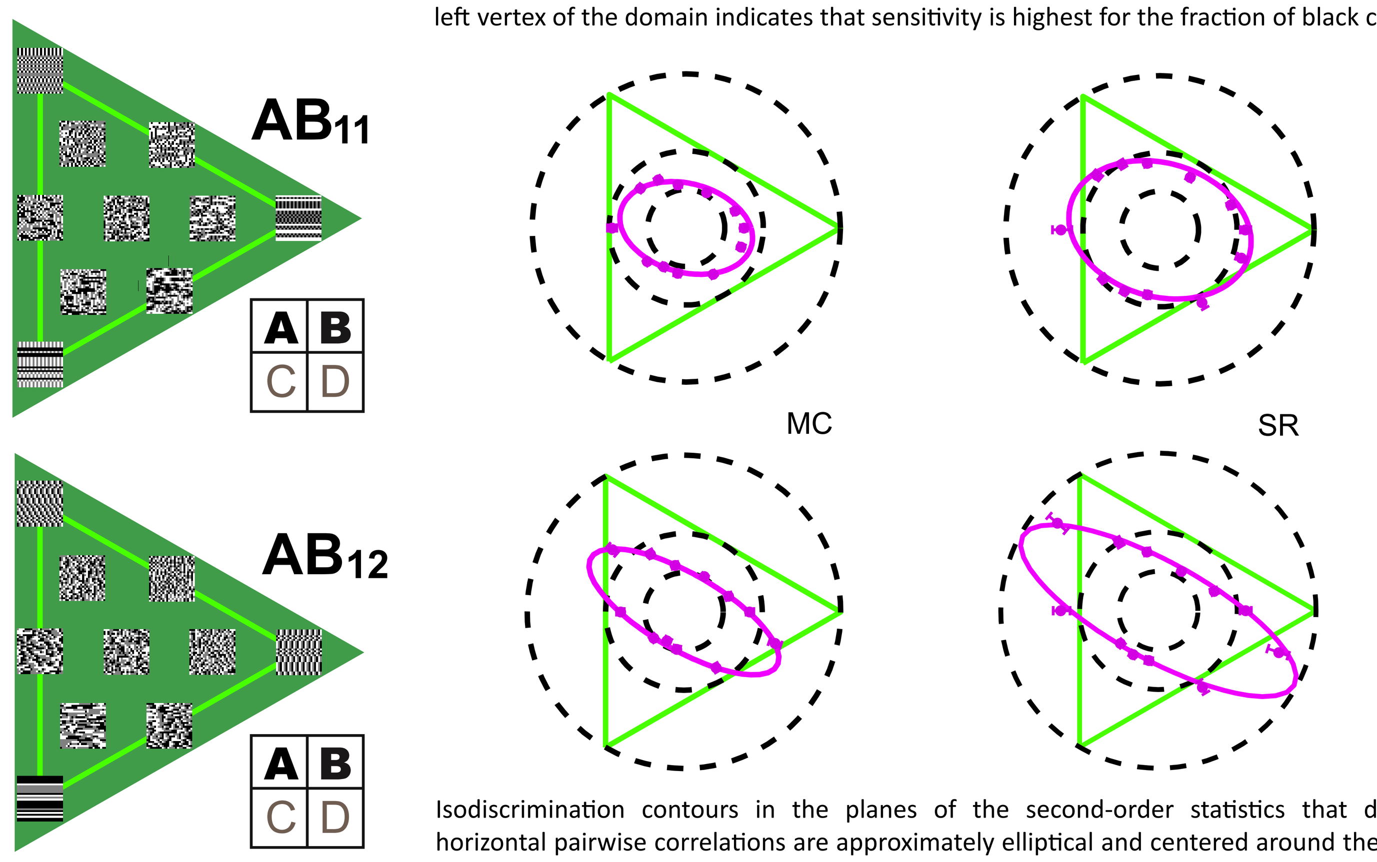
Victor, J.D., Rizvi, S.M., and Conte, M.M. Simple combination rules for sensitivities to image statistics. Program No. 842.11 2016 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience, 2016. Online.

Victor, J.D., Thengone, D..J., Rizvi, S.M., and Conte, M.M. (2015) A perceptual space of local image statistics. Vision Research **117**, 117-135.

Results: Isodiscrimination Contours

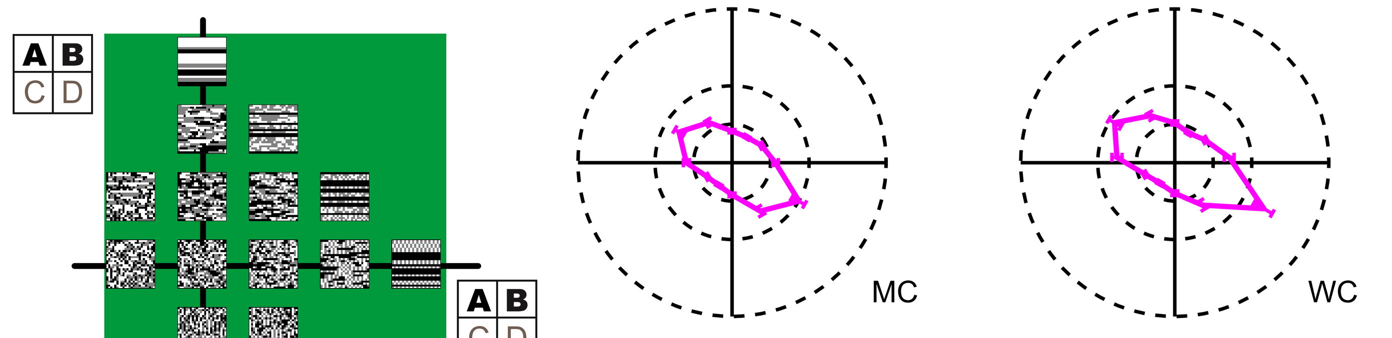


Isodiscrimination contours in the plane of the first-order statistic are approximately elliptical and centered at the origin, indicating similar sensitivities for positive and negative deviations of the image statistics. The orientation of the minor axis of the ellipse towards the lower left vertex of the domain indicates that sensitivity is highest for the fraction of black checks.

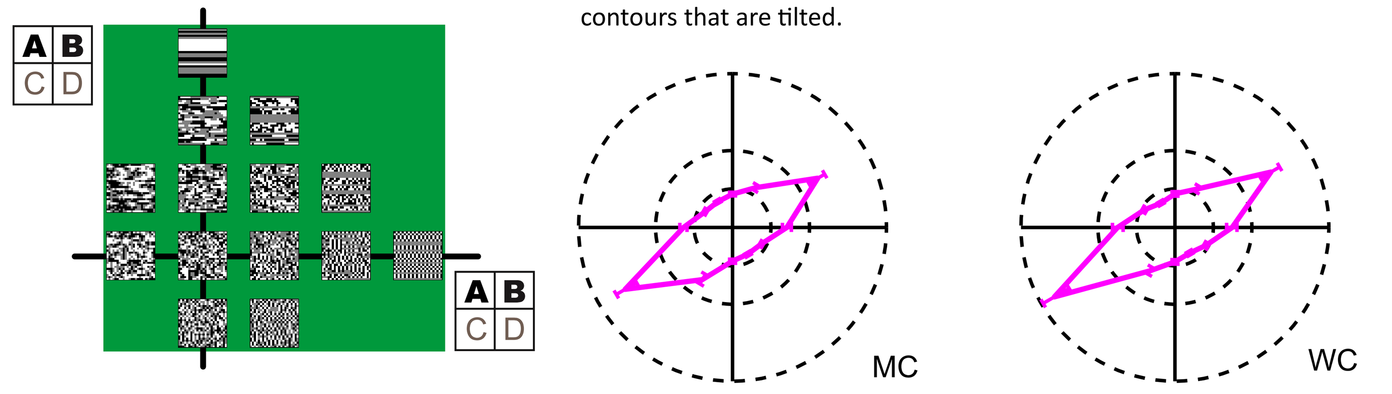


Isodiscrimination contours in the planes of the second-order statistics that describe horizontal pairwise correlations are approximately elliptical and centered around the origin. In the AB_{12}-plane, the ellipse is more eccentric than in the AB_{11}-plane, indicating a lower threshold for correlations that produce horizontal elongated regions than for similar levels of correlations that do not.

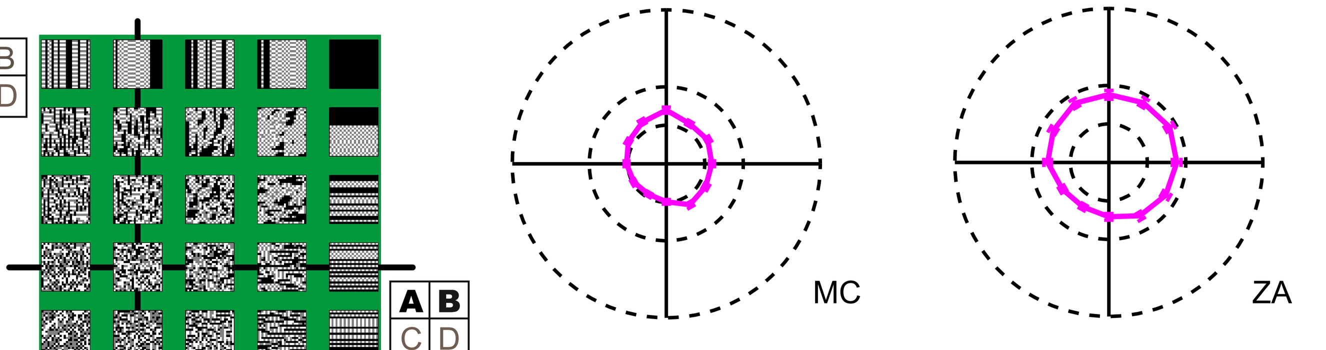
AB_{11} and AB_{12}



When two kinds of horizontal pairwise correlations (AB_{11} and AB_{12}) are present, they interact strongly, leading to elliptical isodiscrimination contours that are tilted.

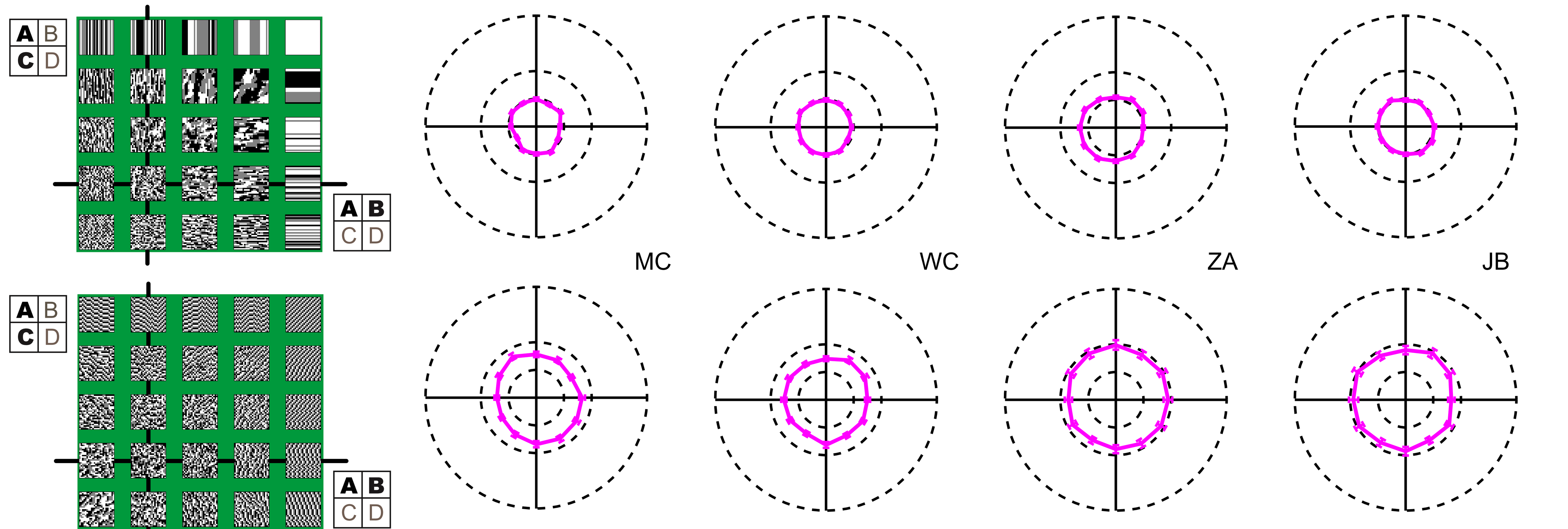


AB_{11} and AC_{11}

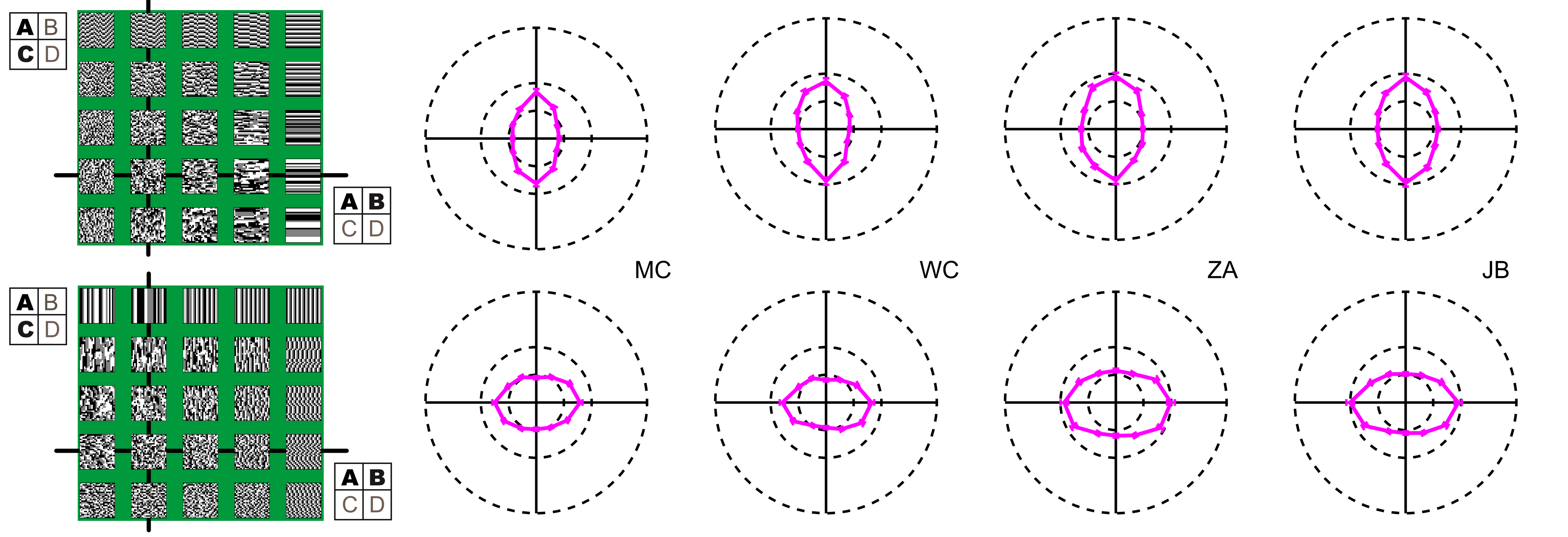


Interactions between horizontal and vertical pairwise correlations that produce single-tone elongated regions (AB_{11} and AC_{11}) are weak. Sensitivities for horizontal and vertical correlations are nearly equal. This leads to nearly circular isodiscrimination contours.

AB_{12} and AC_{12}



For other kinds of horizontal and vertical pairwise correlations (AB_{12} and AC_{12}), sensitivities vary widely. Interactions are weak, leading to circular (above) or elliptical (below) isodiscrimination contours that are aligned with the axes.

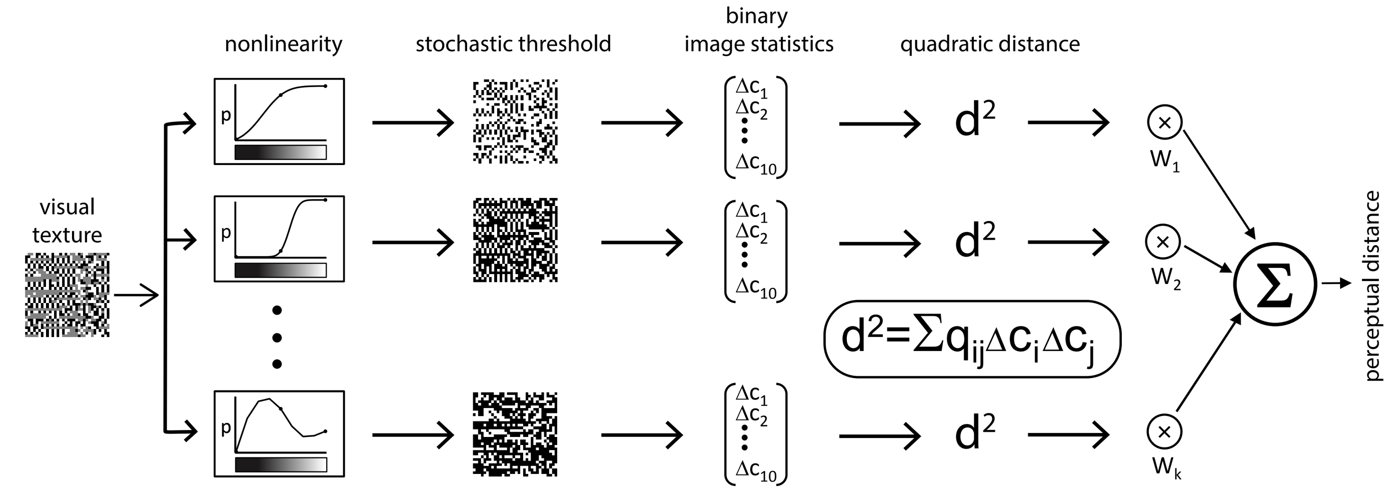


Results Summary

- For textures composed of black, gray and white checks, segmentation can be driven by all first- and second-order statistics (shown here), and selected third- and fourth-order statistics (SFN 2016).
- Segmentation thresholds for increments and decrements of an image statistic are nearly equal.
- Sensitivities combine in a quadratic fashion, yielding elliptical isodiscrimination contours. This holds both for pairs of statistics that lie in the same triangular coordinate domain, and pairs that do not.

A Model Framework

Each channel first transforms the gray-level image into a binary image via a stochastic threshold, in which the probability of assignment to black or white depends nonlinearly on the gray-level value of the original texture. The 2x2 block statistics of these internal representations are computed and compared across regions, yielding 10-element parameter vectors $(\Delta c_1, \Delta c_2, \ldots, \Delta c_{10})$. A quadratic distance is applied to each of these their vectors, and the weighted sum yields the predicted perceptual distance.



Choosing the stochastic nonlinearities according to Chubb et al. (2014) and the quadratic distance parameters q_{ij} according to Victor et al. (2015) yields models consistent with both data sets, and accounts at least qualitatively for the findings reported here.

Conclusions

- The two main findings that emerged from studies of binary textures also apply to textures with gray levels: similar thresholds for positive and negative variations of image statistics, and quadratic cue combination.
- These behaviors are captured by a model that unifies two previous sets of studies: Chubb et al.'s studies of textures with multiple gray levels and no spatial correlations, and our studies of spatially-correlated textures with black and white elements.