One dimensional measures of higher-order image structure predict human performance at discriminating complex form

Ryan R L Taylor1, Ted Maddess2, Yoshinori Nagai3

1,2ARC Centre of Excellence in Vision Science, Australian National University
3Center for Information Science, Kokushikan University, Tokyo, Japan

Higher-order image statistics describe complex and potentially more informative aspects of visual scenes such as corners or intersections. Such features are not accurately described until sets of spatial interactions between four or more points are taken into account [1]. The encoding of such image structure by the visual system cannot be studied with traditional, typically 2nd-order tools such as gratings. Moreover, higher-order spatial correlations are difficult to quantify and to control experimentally. The present research overcomes such difficulties by extending the pioneering texture approach of Julesz [2].

A large battery of textures recently proposed by Maddess and co-workers [3, 4] were employed. The test battery consisted of 21 collections or ensembles each containing individual textures generated by the same set of mathematical rules. Importantly, the average third-order (and lower) correlation functions (3CFs) of each ensemble are not significantly different from zero, as is also the case for evenly distributed noise patterns (pixel values assigned randomly and with equal probability). In order to discriminate such ensembles from each other and from noise patterns one must learn ensemble specific higher-order features. A two alternative forced choice (2AFC) psychophysical study was conducted to determine the average discriminability of each of the 21 ensembles in 24 subjects. Data yielded psychometric functions (PFs) that gave the mean probability of correctly discriminating a given texture ensemble from noise patterns. PFs were compared with higher-order measures of image information (entropy) to gain insight into how subjects may be detecting statistical structure within each of the ensembles. Results suggest that histogram-like entropy statistics based on the occurrence of n-pixel (n > 5) combinations along single or multiple (depending on learning and viewing time) one dimensional horizontal strips within textures were sufficient to accurately capture human performance. The best $R^2$ fit between PFs and horizontal entropy measures was 0.97. These findings may suggest how local filter information may be pooled across space to confer higher-order sensitivity.

Acknowledgments
This work was supported by the ARC Centre of Excellence in Vision Science. Grant no: CE0561903

References


