Local multipoint correlations support categorial classification of objects and backgrounds

Ann M Hermundstad[1,2], Tiberiu Tesileanu[2,3], John J Briguglio[2], Kalina Slavkova[2], Syed M Rizvi[4], Mary M Conte[4], Jonathan D Victor[4], Vijay Balasubramanian[2,3]

[1]Janelia Research Campus; [2]Dept. Physics & Astronomy, University of Pennsylvania; [3]Initiative for the Theoretical Sciences, CUNY Graduate Center; [4]Feil Family Brain and Mind Res. Inst., Weill Cornell Medical College

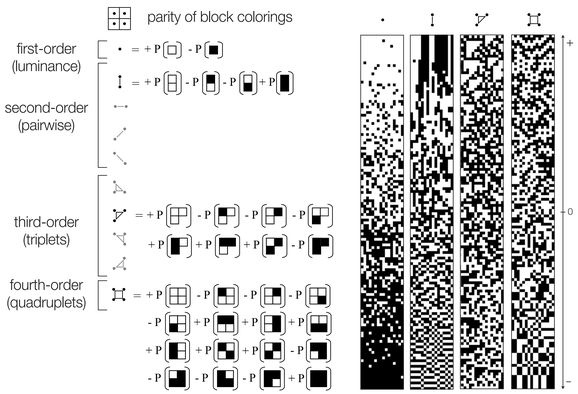
motivation

parsing a visual scene relies on identifying and distinguishing between visual features that capture texture and shape

previous approaches have focused on shape and similar cues; here we consider an independent source of information (texture)

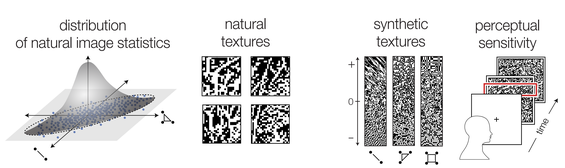
a low-dimensional space of local image statistics can be used to capture visually-salient information about natural textures

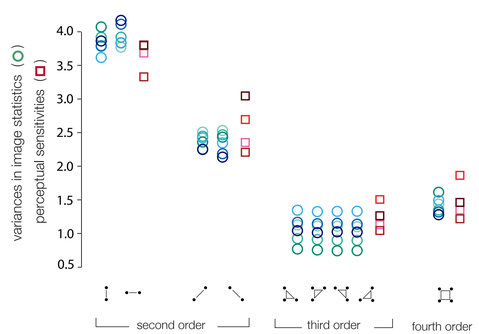
a set of ten independent coordinates captures local multipoint correlations between binarized pixels in an image



these coordinates are informative about the ensemble of natural scenes...

...and the visual system is tuned to represent these coordinates in an efficient manner

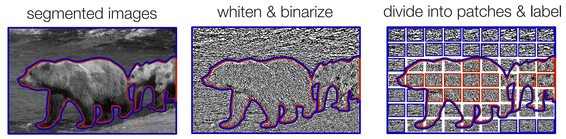




question

to what extent are local image statistics informative about object categorization?

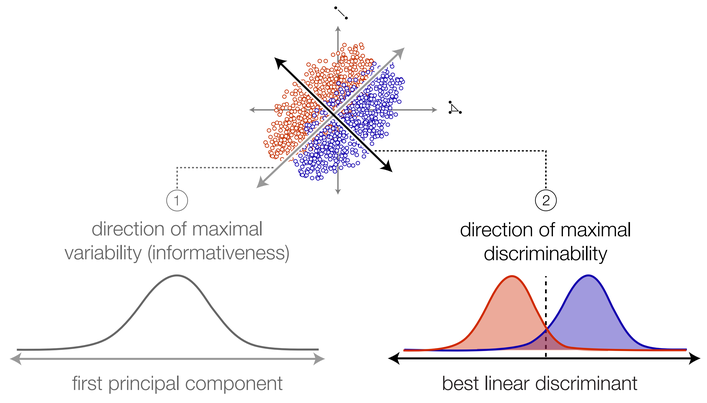
approach



we perform two analyses of local image statistics on different categories of images (right; schematic):

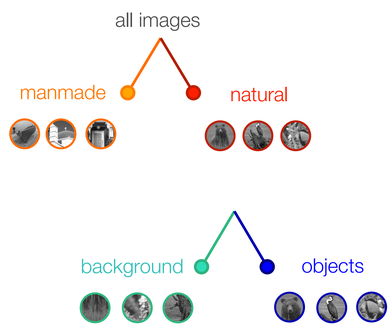
① principal component analysis: direction of maximal informativeness

② linear discriminant analysis: direction of maximal discriminability



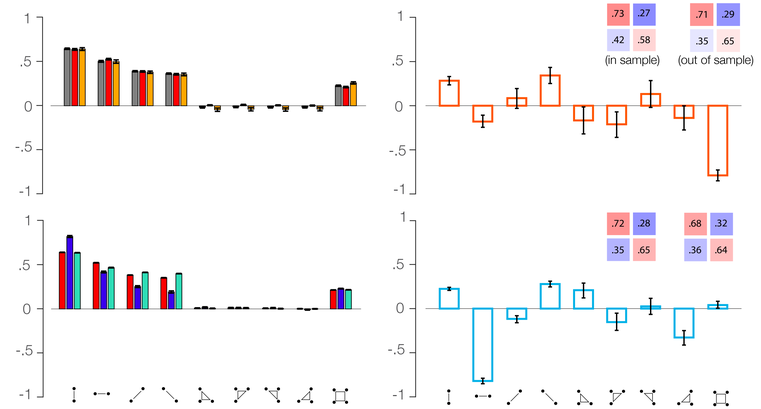
analysis

assign categorial classes to each image patch:



① find first principal component of each class:

② find best linear discriminant between classes:

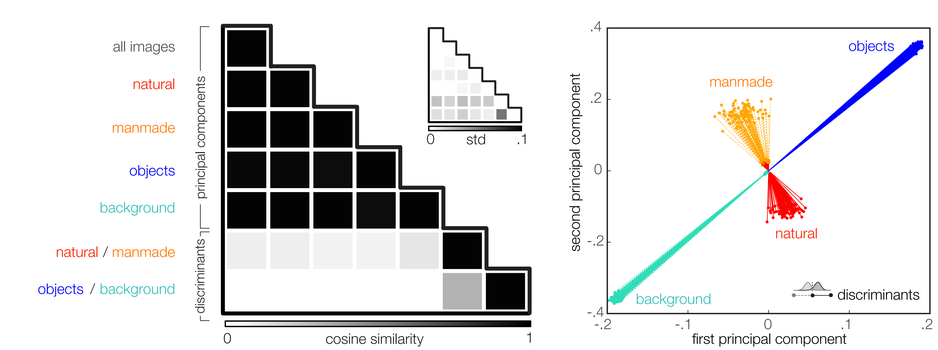


results

directions of maximal informativeness are highly consistent across different image categories

a linear discriminant can separate different categories; classification error is significantly below chance

directions of maximal discriminability involve different combinations of statistics than directions of maximal informativeness



summary

conclusions:

local image statistics are informative about specific aspects of scene content

combinations of image features that are useful for categorization are a subset of those that are useful for segmentation

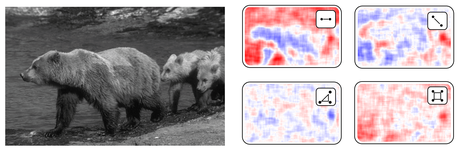
implications:

specific directions in the perceptual space of textures are useful for specific visual tasks

local processing could play an adaptable role in segmentation and categorization

future directions

how does texture information vary with space and scale?



does the visual system exploit different spatial scales for different visual tasks?

does the visual system exploit specific directions in the perceptual space of textures for different visual tasks?

acknowledgements

support:

NIH Grant EY07977; NSF Grant PHY-1058202

AMH was supported by a Burroughs Wellcome Fund Career Award at the Scientific Interface

TT was supported by the Swartz Foundation

references:

[1] Victor JD, Thengone DJ, Rizvi SM, and Conte MM. (2015), Vis. Res. 117, 117-135.

[2] Hermundstad AM, Briguglio JJ, Conte MM, Victor JD, Balasubramanian V, and Tkacik G (2014), eLife, 3, e03722

[3] Victor JD, Conte MM (2012), J. Opt. Soc. Am. A, 29:1313-1345.

[4] Tkacik G, Prentice JS, Victor JD and Balasubramanian V (2010), PNAS,107(42):18149-18154.