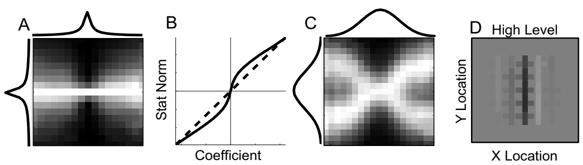
STATISTICAL NORMALIZATION OF SENSORY REPRESENTATIONS

Odelia Schwartz[1] and Terrence J Sejnowski[1] and Peter Dayan[2]
[1]The Salk Institute and HHMI, La Jolla CA and [2]GCNU, UCL, London

A critical question in the neural representations of the sensory world, is how sensory systems can build multi-layer, hierarchical, representations whose semantics are as uniform as possible at every stage. Here, we demonstrate that a *statistically* normalizing transform on wavelet outputs can license effective hierarchical representations. The figure below outlines our technique in the case of images. (A) shows a classic bowtie dependence between two spatially adjacent wavelet coefficients; and (A: top and left) shows the sparse marginal distributions of the coefficients. We suggest performing the non-linear *marginal* transform in (B), giving rise to a standard Gaussian marginal distribution (hence statistically normalized; C: top and left), and a contrast enhanced bowtie (C), with a bimodal conditional distribution of one coefficient given the other. Applying ICA to the non-linearly transformed coefficients, yields higher level units capturing long range structure in the inputs (D; wavelet feedforward weights, black negative; white positive).



We consider statistical normalization in relation to three related frameworks: Gaussian scale mixture models (GSMs), divisive normalization and sparse coding. Crudely, GSMs model sensory input in terms of a hierarchical *scale* process multiplying a hierarchical standard Gaussian. Sparse coding models such as independent component analysis (ICA) gain purchase because of the leptokurtotic properties of GSMs, extracting wavelet-like filters. ICA typically cannot be iterated, since the untransformed coefficients are already strongly sparse. Divisive normalization can eliminate the bowtie dependence arising because of the common scale factors; but unsupervised learning based on the transformed outputs is difficult since they have strong joint Gaussian structure. In our technique, the contrast enhanced bowties arise because the normalizing transform ties the output to an estimate of the local scale factor, and hence to GSMs (we expect optimizing the target marginal in (C) to make this relation more precise). Our results suggest that this allows hierarchical ICA to succeed (D). We are exploring properties of the contrast enhanced bowties for visual and auditory inputs; and connections with adaptation and, via conditional rather than marginal standardization, divisive normalization.