

Tapestries of experts and a novel path to score matching

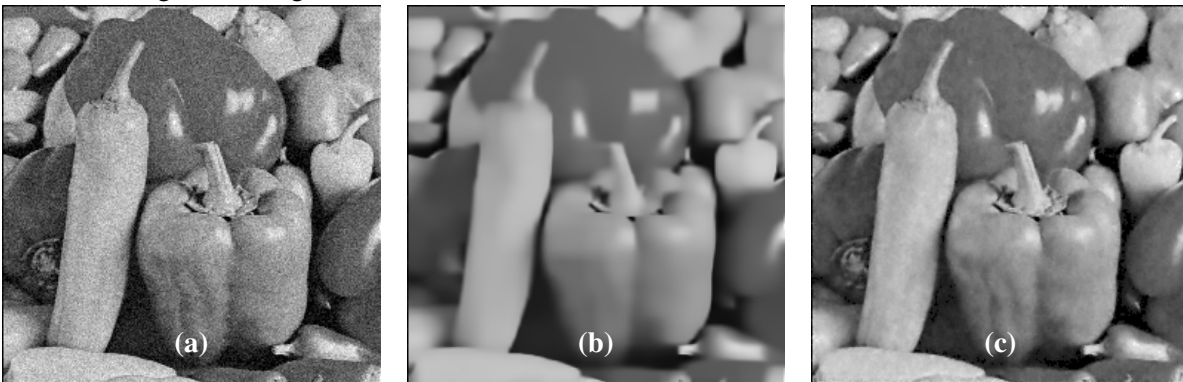
Jascha Sohl-Dickstein, Bruno Olshausen

Olshausen Lab/Redwood Center, University of California, Berkeley

We present the Tapestry of Experts (TOE) model, a spatially heterogeneous generalization of the Field of Experts (FOE) [2] model in which individual experts need not be concerned with interference by spatially shifted versions of themselves. In addition we present an independent derivation of the score matching [1] objective function for performing maximum likelihood estimation on models (like FOE and TOE) which cannot be analytically normalized, and consider extensions suggested by our derivation.

The TOE model we propose has a model distribution $p(\mathbf{x}) = \frac{1}{Z(\theta)} \prod_{i=1}^M \prod_{j=1}^M \prod_{k=1}^O \phi_{(i \bmod N, j \bmod N, k)}(\mathbf{x} | 0 < i_x < N, 0 < j_x < N)$ where N, M are the widths of receptive field and image, O is the degree of overcompleteness and $\phi_i(\mathbf{x}) = \left(1 + \frac{1}{2} (\mathbf{J}_i^T \mathbf{x})^2\right)^{-\alpha_i}$.

As in FOE the experts tile the data space, but unlike in FOE experts only recur when they will no longer overlap themselves. TOE appears better able to capture the statistics of natural scenes - as demonstrated below in an image denoising task:



An image (a) with additive Gaussian noise ($\sigma = 0.3\sigma_{image}$) and MAP estimates of the original image from (b) a 24x overcomplete FOE model with 5x5 receptive fields and (c) a 6x overcomplete TOE model with 5x5 receptive fields. The $Z(\theta)$ normalization term - the so called partition function - is intractable in both FOE and TOE models. We sidestep this difficulty by approximating the learning gradient using an expansion around the datapoints of the difference in energy gradients between data and model distributions. In contrast to traditional methods for dealing with intractable partition functions, such as mean field or renormalization group theory, this technique depends solely on local properties of the model rather than macroscopic or scale invariant properties. The derived objective function $\left(\left\langle \frac{1}{2} \nabla_x E(x; \theta) \cdot \nabla_x E(x; \theta) - \nabla_x \cdot \nabla_x E(x; \theta) \right\rangle_{p(x)}\right)$ turns out to be precisely that derived by Hyvärinen via score matching. The derivation we present places score matching more firmly in a context with other approximate analytical solutions to the partition function problem however, and promises several extensions. It is possible for instance to include higher order terms in the derivation, increasing the size of the patches around the datapoints over which the model distribution can be approximated.

References

- [1] Estimation of non-normalized statistical models using score matching. A. Hyvärinen, *JMLR* 6:695-709, 2005.
- [2] Fields of Experts: a framework for learning image priors. S. Roth and M. Black, *IEEE Proceedings* 860-867, June 2005