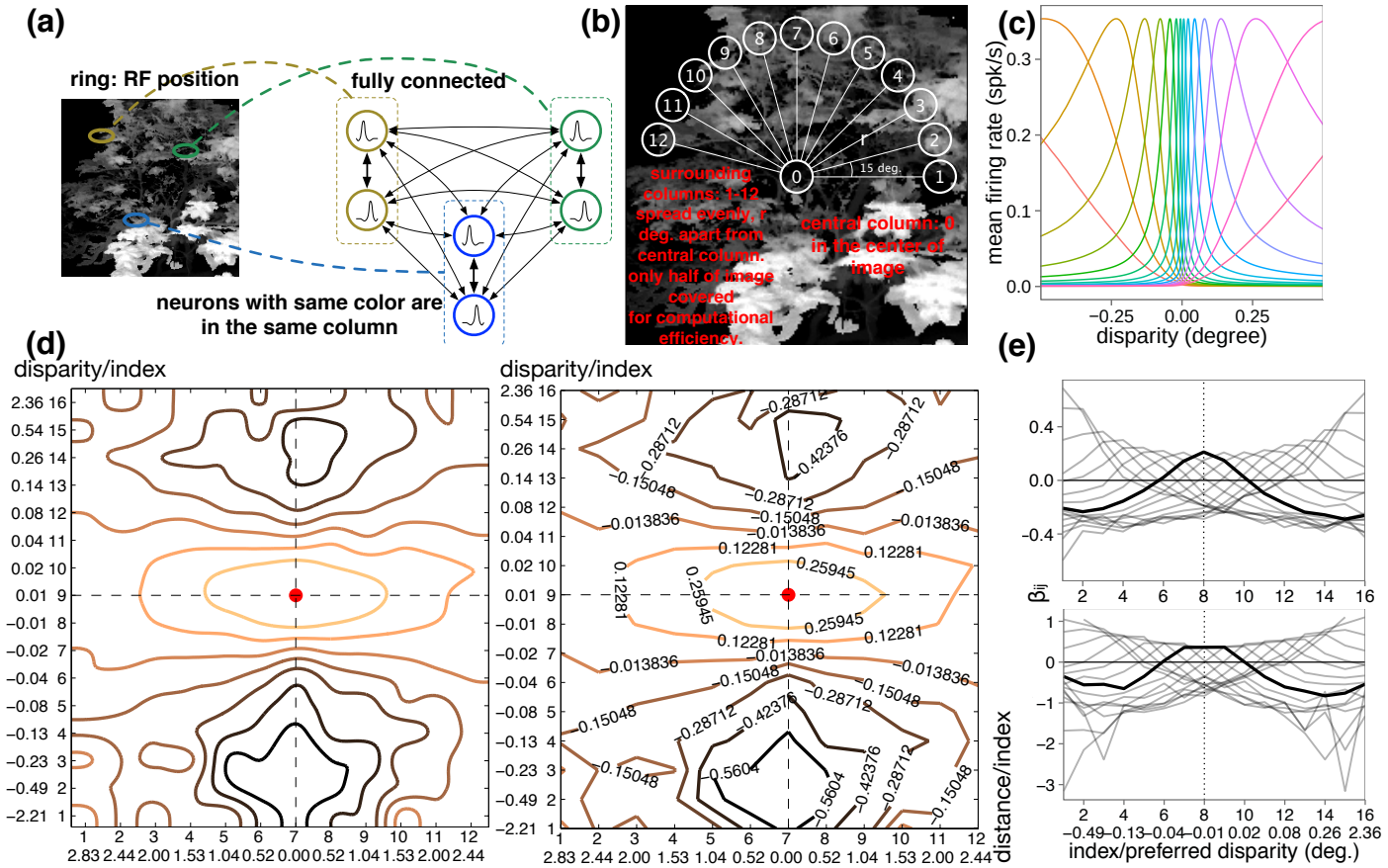
Relating functional connectivity in V1 neural circuits and 3D natural scenes using Boltzmann machines

Authors Yimeng Zhang, Xiong Li, Jason M. Samonds, Ben Poole, Tai Sing Lee

Summary Bayesian theory has provided a compelling conceptualization for perceptual inference in the brain. To understand the neural mechanisms of Bayesian inference, we need to understand the neural representation of statistical regularities in the natural environment. Here, we investigated empirically how the second order statistical regularities in natural 3D scenes are represented in the functional connectivity of a population of disparity-tuned neurons in the primary visual cortex of primates. We applied the Boltzmann machine to learn from 3D natural scenes and found that the functional connectivity between nodes exhibited patterns of cooperative and competitive interactions that are consistent with the observed functional connectivity between disparity-tuned neurons in the macaque primary visual cortex. The positive interactions encode statistical priors about spatial correlations in depth and implement a smoothness constraint. The negative interactions within a hypercolumn and across hypercolumns emerge automatically to reflect the uniqueness constraint found in computational models for stereopsis. These findings demonstrate that there is a relationship between the functional connectivity observed in the visual cortex and the statistics of natural scenes. This relationship suggests that the functional connectivity between disparity-tuned neurons can be considered as a disparity association field. They also suggest that the Boltzmann machine, or a Markov random field in general, can be a viable model for conceptualizing computations in the visual cortex, and as such, can be used to leverage the natural scene statistics to understand neural circuits in the visual cortex.

Methods We used a Boltzmann machine (Hinton & Sejnowski 1986), a.k.a. Ising model (Schneidman et al. 2006), to model the disparity signals across space in 1.1 million disparity image patches generated from 3D range data (Potetz & Lee 2003; Huang et al. 2000), and to compare the connectivity learned by the model with the functional connectivity inferred from neurophysiological data (Samonds et al. 2009). The model network consists of 13 "hypercolumns", arranged spatially in a fan-like structure (one central and twelve surrounding ones), each with a bank of 16 disparity-tuned neurons, as shown in Figures 1a (fewer neurons shown for clarity) and 1b. The neurons at each hypercolumn have receptive fields centered at a specific location within each disparity input data patch. We used equations in Ganguli & Simoncelli (2010) to generate a set of tuning curves for our model neurons in each hypercolumn from the disparity distribution of scene data, as shown in Figure 1c. A Boltzmann machine model was fitted to explain the spiking patterns of $N = 208$ model neurons in response to the disparity image patches. From each patch, disparities at 13 locations shown in Figure 1b were translated into a sequence of spiking pattern $\boldsymbol{x} = [x_1, \ldots, x_N]^T \in \{0, 1\}^N$, with $x_i = 1$ indicating that the i-th neuron has spiked and $x_i = 0$ otherwise at a particular time bin, based on tuning curves and Poisson processes. $\boldsymbol{x}$ is modeled by the Boltzmann machine as $P(\boldsymbol{x}; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{Z(\boldsymbol{\alpha},\boldsymbol{\beta})} \exp\left(\sum_{i=1}^{N} \alpha_i x_i + \sum_{i<j} \beta_{i,j} x_i x_j\right)$, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ model the first and second order interactions among neurons, and $Z(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is the partition function. The parameters of the Boltzmann machine were learned to explain the observed sequence of spike patterns, using minimum probability flow learning algorithm (Sohl-Dickstein et al. 2011). After training, to compare with neurophysiological measurements, we used a conditional random field (or stimulus-dependent max-entropy models (Granot-Atedgi et al. 2013)) to introduce input stimuli drive to the model.

Results Figure 1d shows part of the network connections (for neuron 9 in the 7th surrounding hypercolumn, see Figure 1b) of the trained Boltzmann machine. The right panel of Figure 1b is the contour map computed from the raw connection matrix with labels and the left one is a smoothed version or clarity. We call this the **disparity association field**, analogous to the well-known contour association field. The ordinate indicates neurons with different disparity tunings (tick marks show neurons' indices/preferred disparities) and the abscissa indicates different hypercolumns or spatial locations (tick marks show neurons' indices/distances to column 7). This association field has a number of characteristics. First, in terms of **inter-columnar connections**, neurons with the same and similar disparity tunings tend to form positive connections across hypercolumns and neurons with very different disparity tunings form negative connections. Second, the strength of the connections drops off with distance. Figure 1e (top) shows in more detail how each neuron in one hypercolumn is connected to neurons of various disparity tunings in the adjacent hypercolumns. The dark bold curve highlights the connectivity of neuron 8 in the central hypercolumn (the curve's value at, say, neuron = 9 is the average



of connections from neuron 8 in the central column to every neuron 9 in surrounding columns, and so on), matching the continuity constraint of stereopsis in Marr & Poggio (1976). The **intra-columnar interactions** (Figure 1e bottom, showing connectivity in the same column, with central column chosen for illustration) exhibited excitation for very similarly tuned neurons in the same hypercolumn, but exerted a progressively suppressive effect on neurons of dis-similar disparity tuning, implementing what Marr & Poggio (1976) called the uniqueness constraint. This is almost exactly the same interaction profile specified in Samonds et al. (2013) to account for the dynamics and behaviors of the disparity-tuned neurons in our earlier neurophysiological data.

We also compared more qualitatively the functional connectivity in neurophysiological data measured from CCHs with that of the augmented Boltzmann machine model or conditional random field. We found matches between them in at least 3 aspects: 1) They both decay over distance between neurons (Smith & Kohn 2008). 2) The connectivity strength is proportional to similarity in disparity tuning (Samonds et al. 2009). 3) Cardinal effect: there's stronger connectivity between neurons in vertical & horizontal configurations compared to those in oblique configurations.

Discussion Our findings suggest that certain aspects of the cortical circuits can be predicted from natural scenes using a Boltzmann machine, allowing us to leverage structures in natural scenes to understand structures in neural circuits. Also, the cortical circuit among disparity-tuned neurons appears to form a disparity association field that could be useful for surface interpolation. Finally, the Boltzmann machine or its related class of models might be a viable computational framework to reason about statistical inference in the visual cortex. Boltzmann machine's units can have excitatory and inhibitory connections to other neurons, while real neurons are either excitatory or inhibitory. The negative connections can be implemented through global surround suppression or normalization mechanisms, and a new type of Boltzmann machine will be developed to learn such structures automatically.

References: • Ganguli, D. & Simoncelli, E. P. 2010, in NIPS 23. • Granot-Atedgi, E. et al. 2013, PLoS Comput Biol. • Hinton, G. E. & Sejnowski, T. J. 1986, in PDP. • Huang, J., Lee, A. B., & Mumford, D. 2000, in CVPR 2000. • Marr, D. & Poggio, T. 1976, Science. • Potetz, B. & Lee, T. S. 2003, JOSA A. • Samonds, J. M., Potetz, B. R., & Lee, T. S. 2009, J. Neurosci. • Samonds, J. M. et al. 2013, J. Neurosci. • Schneidman, E. et al. 2006, Nature. • Smith, M. A. & Kohn, A. 2008, J. Neurosci., 28, 12591. • Sohl-Dickstein, J., Battaglino, P., & DeWeese, M. 2011, in ICML 2011.