



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



Yimeng Zhang<sup>a</sup>, Xiong Li<sup>a,b</sup>, Jason M. Samonds<sup>a</sup>, Tai Sing Lee<sup>a,\*</sup>

<sup>a</sup> Center for the Neural Basis of Cognition and Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA

<sup>b</sup> Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China

## ARTICLE INFO

### Article history:

Received 8 September 2014

Received in revised form 3 December 2015

Accepted 7 December 2015

Available online 28 December 2015

### Keywords:

Visual cortex

Association field

Binocular disparity

Scene statistics

Functional connectivity

## ABSTRACT

Bayesian theory has provided a compelling conceptualization for perceptual inference in the brain. Central to Bayesian inference is the notion of statistical priors. To understand the neural mechanisms of Bayesian inference, we need to understand the neural representation of statistical regularities in the natural environment. In this paper, we investigated empirically how statistical regularities in natural 3D scenes are represented in the functional connectivity of disparity-tuned neurons in the primary visual cortex of primates. We applied a Boltzmann machine model to learn from 3D natural scenes, and found that the units in the model exhibited cooperative and competitive interactions, forming a “disparity association field”, analogous to the contour association field. The cooperative and competitive interactions in the disparity association field are consistent with constraints of computational models for stereo matching. In addition, we simulated neurophysiological experiments on the model, and found the results to be consistent with neurophysiological data in terms of the functional connectivity measurements between disparity-tuned neurons in the macaque primary visual cortex. These findings demonstrate that there is a relationship between the functional connectivity observed in the visual cortex and the statistics of natural scenes. They also suggest that the Boltzmann machine can be a viable model for conceptualizing computations in the visual cortex and, as such, can be used to predict neural circuits in the visual cortex from natural scene statistics.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Natural scenes contain significant ambiguity. To resolve ambiguities and obtain a stable 3D percept of the world, the visual system (as well as the whole brain) must perform inference, integrating current sensory data with prior knowledge of the world formulated from past experience. Therefore, (Bayesian) inference has long been proposed as a fundamental computational principle of the brain (von Helmholtz, 1896; Knil & Richards, 1996). In this work, we attempt to address one of the key questions for understanding Bayesian inference in the brain, in the context of the primary visual cortex (V1): how might an internal model of natural scenes—the Bayesian prior—be encoded in the brain?

To support visual inference, an internal representation of the visual scenes requires encoding both the statistical regularities of the boundaries and of the surfaces themselves. There have been studies suggesting that the neural circuits in the primary visual

cortex (V1) encode contour priors in the form of the contour association field (Field, Hayes, & Hess, 1993; Kapadia, Ito, Gilbert, & Westheimer, 1995; Geisler, Perry, Super, & Gallogly, 2001; Elder & Goldberg, 2002; Bosking, Crowley, & Fitzpatrick, 2002; Li & Gilbert, 2002; Menz & Freeman, 2003; Samonds, Potetz, & Lee, 2009; Samonds, Potetz, Tyler, & Lee, 2013). Recent neurophysiological evidence suggests that disparity-tuned neurons in the primary visual cortex might form a recurrent network for stereo processing (Samonds et al., 2009; Samonds et al., 2013). This network encodes the statistical correlation of disparity signals in natural scenes, complementing the contour association field, and is referred to as the disparity association field. However, the neural mechanisms by which statistical priors of boundaries and surfaces from the environment can be learned are not well understood.

We hypothesize that the empirically observed neural connectivity between disparity-tuned neurons in V1 can be predicted from 3D natural scenes using a Boltzmann machine. To test this hypothesis, we fitted a Boltzmann machine neural network model (Hinton & Sejnowski, 1986) to disparity signals derived from 3D natural scene data, and found that (1) learned parameters in the model were consistent with connectivity constraints in stereopsis

\* Corresponding author.

E-mail addresses: [yimengzh@cs.cmu.edu](mailto:yimengzh@cs.cmu.edu) (Y. Zhang), [flit.lee@gmail.com](mailto:flit.lee@gmail.com) (X. Li), [samondjm@cnbc.cmu.edu](mailto:samondjm@cnbc.cmu.edu) (J.M. Samonds), [tai@cnbc.cmu.edu](mailto:tai@cnbc.cmu.edu) (T.S. Lee).

models (Marr & Poggio, 1976; Samonds et al., 2013); (2) the model was consistent with neurophysiological data in terms of functional connectivities among disparity-tuned neurons in V1 (Samonds et al., 2009). The results provide further evidence in support of the notion of the disparity association field, and demonstrate that the Boltzmann machine is a viable model for describing cortical computation in the sense that they can be used to predict functional neural circuitry in the visual cortex.

The paper is organized as follows. In Section 2, we describe the 3D natural scene data and the Boltzmann machine model, as well as the neurophysiological experiments for measuring functional connectivities between pairs of neurons. In Section 3, we compare the trained Boltzmann machine with computational models and neurophysiological data. In Section 4, we discuss the potential implications of this model and its limitations.

## 2. Methods

### 2.1. 3D scene data

We trained a Boltzmann machine to model the disparity signals over a small visual field. These signals were derived from the Brown Range Image Database (Huang, Lee, & Mumford, 2000). A total of 200 K disparity image patches with a  $2^\circ$  half-width were extracted from 172 images (54 forest, 49 interior, 69 residential). The images in the Brown data set were captured by a scanner with range at 2–200 m, and image resolutions were approximately 5 pixels per degree of visual angle.

Disparity image patches were generated from each range image as follows (Fig. 1b). A random point in the range image was chosen as the fixation point. Given the fixation point, the disparities at its surrounding pixels were computed using the method in Liu, Bovik, and Cormack (2008) (see Section 2.1.1 for detail). Finally, a disparity image patch with a  $2^\circ$  half-width was extracted  $3^\circ$  away from the fixation point. This eccentricity was chosen to roughly match the typical receptive field locations of recorded V1 neurons in our earlier neurophysiological experiments.

#### 2.1.1. Disparity computation

We used an optical model of the primate eye following Liu et al. (2008) to compute disparity. In this model (Fig. 1a), each eye is approximated as a perfect sphere centered at its nodal point, and inter-pupillary distance is assumed to be 0.038 m with nodal points at  $(-0.019, 0, 0)$  and  $(0.019, 0, 0)$  as in monkey physiology.

Consider some fixation point  $F = (x_f, y_f, z_f)$ . Let  $O_c = (0, 0, 0)$  be the midpoint between the two eyes. We assume all observations are directed along the  $-z$  axis, or  $x_f = y_f = 0$ . The distance from  $O_c$  to  $F$  is then just  $z_f$ . The horizontal disparity,  $d$ , of an arbitrary point  $P = (x_p, y_p, z_p)$ , is given by

$$d = \beta_r - \beta_l = \alpha - \phi, \quad (1)$$

$$\alpha = 2 \operatorname{atan}(-0.019/z_f), \quad (2)$$

$$\phi = \operatorname{atan}\left(\frac{-x_p - 0.019}{z_p}\right) - \operatorname{atan}\left(\frac{-x_p + 0.019}{z_p}\right). \quad (3)$$

We made the simplifying assumption that fixations occur at any point in the scene with uniform probability. This assumption is supported by Liu et al. (2008), which shows that random fixations roughly emulate the statistics of fixation, at least in natural scenes. This assumption should not affect the basic conclusion of our results.

### 2.2. Boltzmann machines

#### 2.2.1. Interaction among neurons modeled by Boltzmann machines

The extracted disparity image patches reflect the prior of disparity signals in the natural scene, and we modeled this prior by fitting a Boltzmann machine to the patches. Boltzmann machines (Hinton & Sejnowski, 1986) are a class of stochastic recurrent neural network models that can learn internal representations to explain or generate the distribution of the input data, using pairwise connectivity between units to encode the structures of the input data. Boltzmann machines are also a type of Markov random fields, which are widely used in computer vision for solving a variety of early vision problems such as surface interpolation and stereo inference (Geman & Geman, 1984; Koch, Marroquin, & Yuille, 1986; Belhumeur & Mumford, 1992; Tappen & Freeman, 2003). We hypothesize that Boltzmann machines are a viable computational model for understanding the circuitry of the visual cortex, and we will examine if they can explain interactions among neurons in other computational and neurophysiological studies (Marr & Poggio, 1976; Samonds et al., 2009; Samonds et al., 2013). Specifically, the interaction terms  $\beta$  (Eq. (4)) in our Boltzmann machine model were compared with existing computational models in Section 3.2, and neurophysiological experiments were simulated on the model (Section 2.2.2) to compare it with neural data in Section 3.3.

The units in our Boltzmann machine model (Fig. 2a) are organized into a hidden layer and a visible layer, arranged in a spatial 5 by 5 grid of “hypercolumns” (in total  $C = 25$  columns). Each hypercolumn has a bank of  $M = 16$  visible units that encode the disparity input, and a bank of 16 corresponding hidden units  $\mathbf{h}$ , all sharing the same spatial receptive field location. The  $N = MC = 400$  hidden units are fully connected, each of them driven by its corresponding input visible unit. The collective spiking activity at each bank of visible units encodes the disparity signal at the corresponding hypercolumn.

This model is formally expressed as a probability distribution over hidden and visible units:

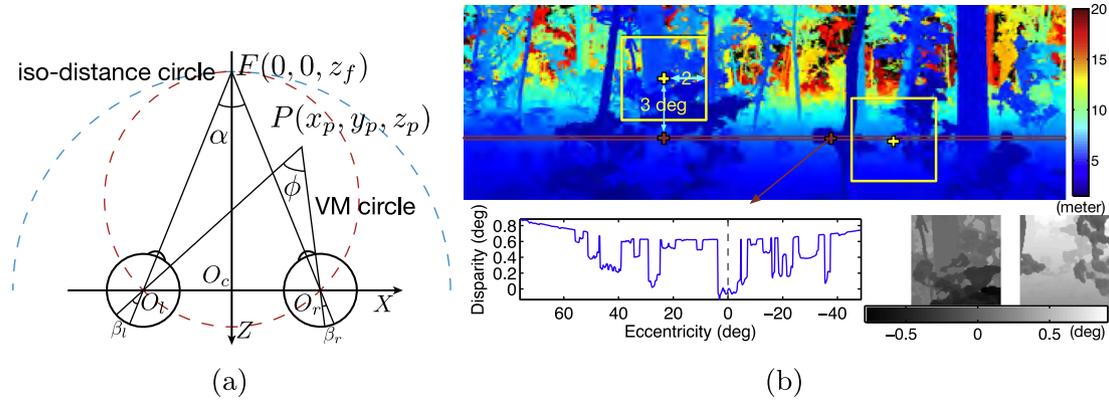
$$P(\mathbf{h}, \mathbf{v}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) = \frac{1}{Z} \exp\left(\sum_{i=1}^N \alpha_i h_i + \sum_{i < j} \beta_{ij} h_i h_j + \sum_{i=1}^N \lambda_i h_i v_i + \sum_{i=1}^N \gamma_i v_i\right). \quad (4)$$

In Eq. (4),  $\mathbf{h}$  and  $\mathbf{v}$  are binary vectors whose distributions are to be captured by the model, representing spiking activities of hidden and visible units. The other model parameters capture the distributions of  $\mathbf{h}$  and  $\mathbf{v}$ , as well as their interactions. Specifically,  $\boldsymbol{\alpha}$  and  $\boldsymbol{\gamma}$  capture the baseline firing rates of hidden and visible units,  $\boldsymbol{\beta}$  models the pairwise lateral interactions among hidden units, and  $\boldsymbol{\lambda}$  models the interactions between hidden and visible units.  $Z$  is a normalization constant.

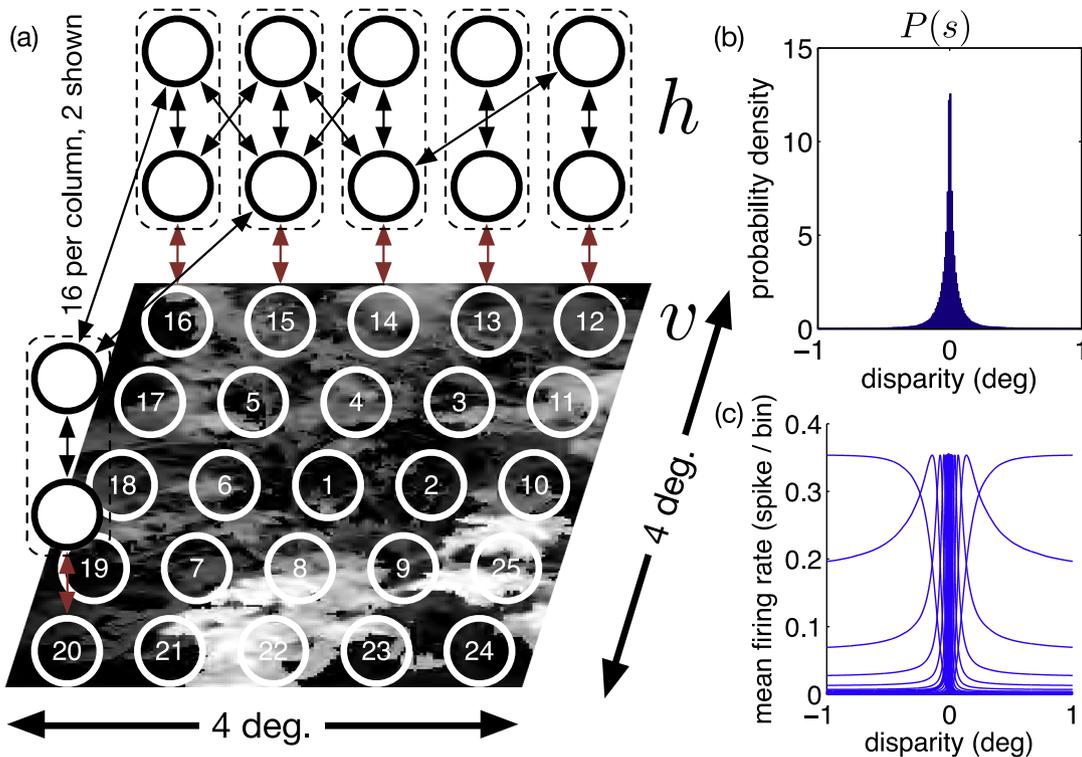
This Boltzmann machine was fitted by finding parameters  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\beta}$ ,  $\boldsymbol{\gamma}$ , and  $\boldsymbol{\lambda}$  that maximize the probability of the model for generating the spike patterns  $\mathbf{v}$ , corresponding to the disparity signals in the extracted patches. Formally, the following log likelihood was maximized:

$$\begin{aligned} L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) &= \sum_{i=1}^T \log P(\mathbf{v}^{(i)}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda}) \\ &= \sum_{i=1}^T \log \sum_{\mathbf{h}^{(j)}} P(\mathbf{h}^{(j)}, \mathbf{v}^{(i)}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\lambda}). \end{aligned} \quad (5)$$

In Eq. (5),  $\mathbf{v}^{(i)}$ 's are  $T$  binary spike patterns of the visible units converted from the disparity signals based on the tuning curves of the visible units (see Fig. 3 and Section 2.2.1.1). The likelihood of observing  $\mathbf{v}^{(i)}$  is computed as the sum of  $P(\mathbf{h}^{(j)}, \mathbf{v}^{(i)})$  over all pos-



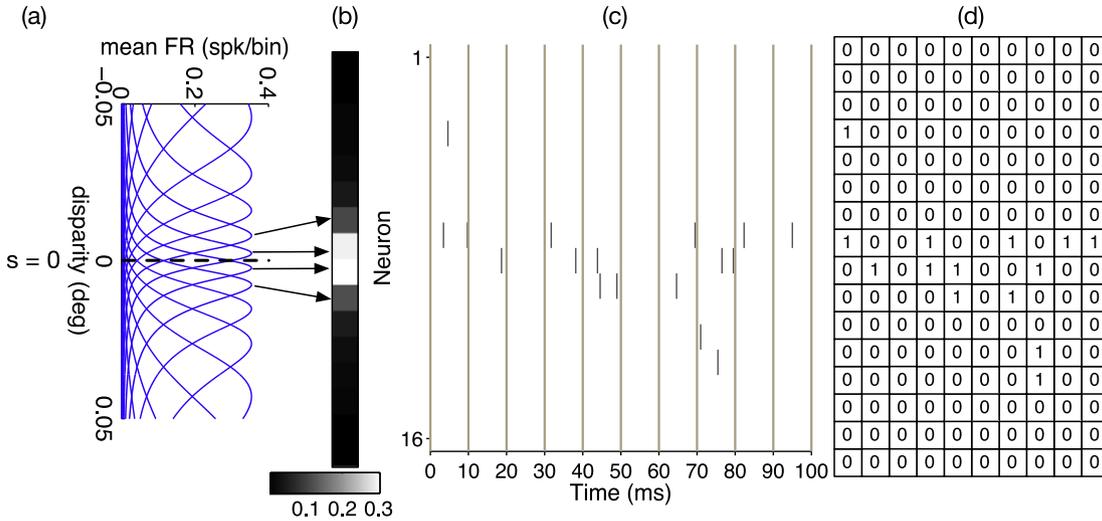
**Fig. 1.** (a) Diagram for calculating disparity. Adapted from Liu et al. (2008). See Eqs. (1)–(3) for detail. (b) One sample range image from the Brown data set (upper) with disparity values along one line in it (lower left), and two extracted disparity patches (lower right). In the upper image: red crosses, fixations points for two patches; yellow crosses, center of patches; red long rectangle, the row shown disparities. Patches were 3° away from fixation and had a half-width of 2°.



**Fig. 2.** Schematic of our Boltzmann machine model (a), distribution of extracted disparity values  $P(s)$  (b), and derived tuning curves of input visible units, with one curve highlighted (c). (a) 25 “hypercolumns” laid in a 5 by 5 grid covering a 4° by 4° patch, with hidden units ( $h$ , black outline) in the same column grouped in dashed box. Each hidden unit has connections (black) to all other ones, and one connection (red) to its own visible unit ( $v$ , white outline). At most two hidden units and one visible unit drawn per column, with many connections missing for clarity. Columns are numbered for reference in Section 3. (b) The distribution of extracted disparity values was sharp, and peaked at zero. (c) Tuning curves of  $v$  were derived based on Ganguli and Simoncelli (2010) with the following details: “baseline” curve was a  $t$  distribution with d.o.f.  $\nu = 2$ , total expected firing rate ( $R$  in Ganguli and Simoncelli (2010)) was unity, and “infomax” optimization criterion was used. Only tuning curves between  $-1^\circ$  and  $1^\circ$  are shown for clarity. Given the sharp distribution of disparity values, the theory in Ganguli and Simoncelli (2010) made the tuning curves at large disparities different from those close to zero. Instead of Gaussian-like (see Fig. 3a for a zoom-in view of tuning curves close to zero), the tuning curves at the two ends of the input distribution were relatively flat at large (positive/negative) disparities, and dropped to zero near zero disparity. Interestingly, these were very similar to the near-tuned and far-tuned neurons in Poggio and Fischer (1977) and Poggio et al., 1988. We also tried Gaussian distribution as the “baseline” curve, but that gave much sharper tuning curves and less co-activation between dissimilar units, which resulted in a less biologically realistic training result.

sible  $2^N$  hidden unit patterns, which is the marginal probability for the model to generate  $\mathbf{v}^{(i)}$ , regardless of the hidden units. Finally, the log probability for the model to generate all the input spike patterns due to the disparity signals is computed as the sum of log probabilities for generating each particular spike pattern  $\mathbf{v}^{(i)}$ . The model was trained using contrastive divergence mean field learning (Welling & Hinton, 2002). See Section 2.2.3 and Welling and Hinton (2002) for more detail.

**2.2.1.1. Conversion of disparity signals into binary spike patterns.** From each disparity image patch  $i$ , disparity values  $s_1^i, s_2^i, \dots, s_{C=25}^i$  corresponding to the locations of the 25 hypercolumns were extracted, and the model was fitted to explain these disparity values across all patches. Disparity signals are real-valued, and must be converted into binary spike patterns, which can be considered as the spiking activities of the bottom-up input to V1 neurons. Following the approach of Ganguli and Simoncelli



**Fig. 3.** Generation of training data for one disparity value. Given one disparity value (a) (in this case  $s = 0$ ), we transformed it into  $M = 16$  mean firing rates (b) using tuning curves (between (a) and (b)), generated spike trains (c), and binned it into a binary matrix (d) as training data to the Boltzmann machine.

(2010), we derived a set of  $M = 16$  tuning curves for visible units (same for all the hypercolumns, Fig. 2c) according to the distribution  $P(s)$  of extracted disparity values from all patches (Fig. 2b). Each disparity value was converted to the mean firing rates of  $M = 16$  visible units based on their tuning curves.

Given the above derived tuning curves, for each patch, we first converted the  $C = 25$  disparity values into the mean firing rates of the  $N = 400$  visible units. Then for each of these  $N$  units, a spike train of 200 ms was generated based on its mean firing rate using an independent homogeneous Poisson process, and the whole spike train was partitioned into 20 bins of 10 ms<sup>1</sup>. A “1” was assigned to a bin of a unit if there were one or more spikes for that unit within that time bin; otherwise, a “0” was assigned. The whole generation process (for one disparity value) is schematically shown in Fig. 3.

### 2.2.2. Simulation of neurophysiological experiments on the model

With the trained Boltzmann machine, we can simulate the neurophysiological experiments by providing the visible units  $\mathbf{v}$  with specific experimental stimuli (Section 2.3.1), and collecting the model response as binary spiking patterns of hidden units  $\mathbf{h}$ . Because a Boltzmann machine models the joint distribution of all hidden and visible units, we can compute the model response by sampling from the conditional distribution of hidden units given the visible units:

$$P(\mathbf{h}|\mathbf{v}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma, \lambda) = \frac{1}{Z} \exp \left( \sum_{i=1}^N (\alpha_i + \lambda_i v_i) h_i + \sum_{i<j} \beta_{ij} h_i h_j \right). \quad (6)$$

After generating hidden unit activities  $\mathbf{h}$  by drawing samples from Eq. (6) (see Section 2.2.1 for detail), we compared  $\mathbf{h}$  with neural data (Section 3.3), in terms of functional connectivity using methods described in Section 2.3.2.

**2.2.2.1. Sampling of hidden unit activities given disparity stimulus.** Given the (real-valued) disparity values of the stimulus at  $C = 25$  hypercolumns, we first converted them into mean firing rates for all input visible units according to the tuning curves. Then

<sup>1</sup> Our implementation was written in terms of bins, with no notion of the physical duration of each bin. We arbitrarily assumed each bin to be of 10 ms, for easier comparison with neurophysiological data and other studies based on Ising models (a type of Boltzmann machines).

we obtained each sample of  $\mathbf{h}$  (a  $N$ -dimensional binary vector) in the following MCMC fashion (Koller & Friedman, 2009).

1. Generate a  $\mathbf{v}$  from the Poisson process described in Section 2.2.1.1, based on mean firing rates of visible units.
2. Initialize  $\mathbf{h}$  randomly, run Gibbs sampling for one step<sup>2</sup> based on Eq. (6).
3. Collect the current  $\mathbf{h}$  as a sample.
4. Start over from (1), but when in (2), initialize  $\mathbf{h}$  with the previous sample.

20,000 samples were generated for each stimulus, and every contiguous 100 samples were regarded as the model response in a trial, with 100 samples between trials, resulting in 100 trials of 100 samples per stimulus<sup>3</sup>. In addition, before collecting the first sample, we performed an additional 100 Gibbs sampling steps, as “burn-in”.

### 2.2.3. More implementation details

For results shown in Section 3, 48,960 disparity patches<sup>4</sup> were extracted from the Brown data set to train the Boltzmann machine model, with distance between nearby hypercolumns set to  $1^\circ$  (Fig. 2a). Data was taken in mini-batches of size 1000, and training took 1000 epochs. A learning rate of  $1 \times 10^{-2}$  was used for learning bias terms  $\boldsymbol{\alpha}$  and  $\boldsymbol{\gamma}$ , and half of that for learning lateral connections  $\boldsymbol{\beta}$ . Multiplicative weight decay of  $1 \times 10^{-2}$  for  $\boldsymbol{\beta}$  multiplied by the learning rate was used, and a momentum factor of 0.5 for first five epochs and one of 0.9 for the rest were employed. Five iterations of mean field updates were used per iteration, with damping parameter set to 0.2. Because  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\gamma}$ , and  $\lambda$  offer too many degrees of freedom for

<sup>2</sup> Here one step of Gibbs sampling is defined as in Koller and Friedman (2009), that is, given the initial state of all hidden units  $h_1, h_2, \dots, h_N$ , we randomly choose one unit  $h_i$ , update it based on  $P(h_i|\mathbf{h}_{-i})$  (where  $-i$  means all but  $i$ ), and do this update sequentially for all  $N$  units. These  $N$  updates are collectively referred to as one step.

<sup>3</sup> Given all 20,000 samples, samples 1–100, 201–300, 401–500, ..., 19,801–19,900 were collected as trials, and samples between them (samples 101–200, 301–400, 501–600, ...) were discarded. This yielded 100 samples per trial, and 100 trials per stimulus.

<sup>4</sup> These patches were from 49 interior images in the Brown data set, with 1000 patches per image. A total of 40 patches were dropped because they had missing range data (thus disparity signals) for some hypercolumns. Other patches from different parts of the data set such as forest scenes were also tried, with empirically similar results, also shown in the Results section.

the model fitting, we fixed  $\lambda$  to be positive, all elements equal to 0.5 during training. This encouraged the resulting  $\alpha$  and  $\gamma$  to be negative, and the hidden units to share the same preferred disparities as their visible input units. If we did not constrain  $\lambda$ , we found that the learned  $\alpha$  and  $\gamma$  were positive, and the tuning curves of hidden units would be inverted from those of visible units, both counter-intuitive.

### 2.3. Neurophysiological experiments

#### 2.3.1. Neurophysiological data

We compared the predictions from the trained Boltzmann machine with observations about the neural circuitry that we have reported in previous papers, based on direct pairwise measurements of neuronal spiking activities (Samonds et al., 2009), and a recurrent neural circuit model that predicts neural responses better than the feed-forward energy model (Samonds et al., 2013).

In these earlier experiments, we analyzed neural data recorded using multi-electrode recording techniques from neurons in the primary visual cortex of three awake behaving macaque monkeys. For monkeys D and F, we used 4–8 tungsten in epoxy or glass microelectrodes (Samonds et al., 2009), while for monkey I we recorded from neurons using a chronically implanted multielectrode array with 96 channels (Samonds, Potetz, & Lee, 2012; Samonds et al., 2013). The experimental protocols for these studies were approved by the Institutional Animal Care and Use Committee of Carnegie Mellon University and in accordance with Public Health Service guidelines for the care and use of laboratory animals.

Stimuli were dynamic random dot stereograms (DRDS) presented for one second per trial while the monkey performed a fixation task. Each DRDS defines a uniform fronto-parallel depth plane (i.e. uniform disparity) inside a 3.5° visual angle aperture window over the receptive fields of the neurons being recorded. These stimuli were standard stimuli used to assess disparity tuning of the neurons and were effective in driving disparity-tuned neurons. The dynamic random dot stereogram (DRDS) was composed of 25-percent black and white dots on a mean gray background with a refresh rate of 12 Hz for dot patterns (monitor refresh rate was 120 Hz) at 11 disparities ( $\pm 0.94^\circ; \pm 0.65^\circ; \pm 282^\circ; \pm 0.188^\circ; \pm 0.094^\circ; 0^\circ$ ). Further details about the neurophysiological experimental procedures are described in our previous works (Samonds et al., 2009; Samonds et al., 2012; Samonds et al., 2013).

#### 2.3.2. Neurophysiological measures of interaction

We measured the functional connectivity between pairs of neurons using a cross-correlation histogram (CCH) measure based on standard methods (Aertsen, Gerstein, Habib, & Palm, 1989; Samonds et al., 2009). The probabilities of joint spike occurrences beyond chance at all possible lag times and all times from stimulus onset were computed by measuring the observed probability of joint occurrences and subtracting the expected joint occurrences, which was the outer product of the peristimulus time histograms:

$$C_{xy}(t_1, t_2) = \langle x(t_1), y(t_2) \rangle - \langle x(t_1) \rangle \langle y(t_2) \rangle, \quad (7)$$

where  $x(t), y(t)$  were the spike trains of the two neurons respectively.

The expected joint occurrences  $\langle x(t_1) \rangle \langle y(t_2) \rangle$  account for the stimulus-related response correlation assuming neurons are independent. They were corrected for trial-to-trial changes in the firing rate to remove potential slow sources of correlation that can lead to apparent fast sources of correlation sometimes referred to as an excitability correction (Gerstein & Kirkland, 2001; Brody, 1999; Ventura, Cai, & Kass, 2005; Samonds et al., 2009). This two-dimensional cross covariance histogram was then normalized

by the square root of the product of the auto-covariance histograms for the two neurons:

$$D_{xy}(t_1, t_2) = \frac{C_{xy}(t_1, t_2)}{\sqrt{C_{xx}(t_1, t_1)C_{yy}(t_2, t_2)}}. \quad (8)$$

This produced a two-dimensional histogram of Pearson's correlation coefficients referred to as the normalized cross-covariance histogram or the normalized joint poststimulus histogram. We average this histogram across diagonals to produce a cross-correlation histogram with respect to lag times between pairs of neurons. We then computed the variance of our estimates by bootstrapping with respect to trials (Efron & Tibshirani, 1994; Samonds et al., 2009; Ventura, Cai, & Kass, 2005). The correlation measurements were the areas under the half-height of the peaks of these cross-correlation histograms for pairs of neurons with a central peak (within 10 ms of 0 lag time) more than three standard deviations above or below the cross-correlation histogram from 175 ms to 375 ms lag time and for pairs of neurons where both neurons had significant disparity tuning (1-way ANOVA,  $p < 0.01$ ) (Samonds et al., 2009). For the neurophysiological results shown in this work (Fig. 8b,d,f), the input stimulus for each pair of neurons was set to the the disparity at which the point-wise product of two neurons' tuning curves was at its maximum (Samonds et al., 2009).

We also computed CCH measures for hidden unit activities generated from the simulation of the trained Boltzmann machine model (Section 2.2.2). To compute CCH for each pair of hidden units in the model, we set the input stimulus (equi-disparity stimulus, with the same disparity value at all hypercolumns) to be at the mean of the preferred disparities of their corresponding visible units<sup>5</sup>. After collecting the samples (see Section 2.2.2.1 for detail), we computed the CCH measure following the method described above<sup>6</sup>, but using only the peak of CCH as the CCH measure of this pair of units, since there was no synaptic delay issue in our Boltzmann machine model, given our simulation method.

## 3. Results

We mainly compared the model with existing computational models in terms of connectivity constraints (Section 3.2), and neurophysiological data in terms of functional connectivities (Section 3.3). The model showed qualitative agreement in both aspects. In the following comparisons, the hidden units correspond to the disparity-tuned V1 neurons, likely realized in terms of the complex cells in the superficial layer of V1 where there are extensive horizontal axonal collaterals forming a recurrent network. The visible units provide the bottom-up input to these V1 complex cells, and they encode disparity signals which in the brain are computed by combining monocular left and right eye simple cells based on phase-shift or position-shift mechanisms (Fleet, Wagner, & Heeger, 1996). The input from visible units, or the corresponding signals in the brain, are assumed to be “decorrelated” across space when stimulus correlation is factored out (Ecker et al., 2010). The prior of natural scenes is assumed to be captured by the lateral connectivity among hidden units in the model or among disparity-tuned V1 neurons in the brain. These intrinsic

<sup>5</sup> We did this because the preferred disparities of hidden units were actually those of their input visible units. See Fig. 4a for detail.

<sup>6</sup> Since our model had  $N = 400$  hidden units, we initially computed the CCH for all  $400 \times 399/2 = 79,800$  pairs. Then we kept pairs whose CCH's peak within 50 ms ( $\pm 5$  bins) of 0 lag time was more than 1.5 standard deviations above or below the CCH from 300 ms to 600 ms ( $\pm 30$  to  $\pm 60$  bins). We further removed pairs where at least one unit responded to the input stimulus with firing rate less than half of its peak response relative to its minimum response over all tested stimuli, resulting in 23,048 pairs shown in Fig. 8.

horizontal connections can give rise to noise correlation and other correlated activities among neurons (Smith & Kohn, 2008; Kelly, Smith, Kass, & Lee, 2010; Cohen & Maunsell, 2009).

### 3.1. First order properties of learned hidden units

Fig. 4 shows typical tuning curves of the hidden units obtained from the model simulation of neurophysiological experiments (Section 2.2.2), and the distribution of bias terms  $\alpha$ ,  $\gamma$ . Hidden units shared the same preferred disparity and the general shape as their corresponding input visible units. The bias terms are negative, indicating that the hidden units tend to fire sparsely.

### 3.2. Comparison with computational models in terms of connectivity constraints

The learned lateral connections  $\beta$  among hidden units form what we call the **disparity association field**, analogous to the well-known contour association field for orientation-tuned neurons (Field et al., 1993). The lateral connectivity, or the disparity association field, observed in the trained Boltzmann machine model is qualitatively in agreement with the cooperative and competitive circuits predicted by Marr and Poggio (1976), and with the recent model of Samonds et al. (2013) which has been successful in more accurately accounting for neurophysiological data of disparity-tuned neurons in V1.

We define the disparity association field of a hidden unit as the set of lateral connections between it and other hidden units. Fig. 5a illustrates the disparity association field of one unit tuned near zero disparity in the center column of the  $5 \times 5$  grid, showing its lateral connections  $\beta$  to all other units in the network along a particular direction in the grid. The  $x$ -axis indicates different hypercolumns or spatial locations, and the  $y$ -axis indicates units with different disparity tunings.

The disparity association field learned by the Boltzmann machine has a number of noteworthy features. First, in terms of **inter-columnar connections**, i.e. connections between a unit with units in other hypercolumns, units with the same or similar disparity tunings tended to form positive connections across hypercolumns (spatial receptive field locations) and units with very different disparity tunings formed negative connections. Fig. 5b and 5c show in greater detail how each unit in one hypercolumn was connected to units of various disparity tunings in other hypercolumns. The dark bold line highlights that unit 8 in one hypercolumn formed positive (excitatory) connections to similarly tuned units (units 6, 7, 8, 9) in the other hypercolumns, and negative (inhibitory) connections to units tuned to very different disparities. Second, in terms of **intra-columnar connections**, i.e. connections between units in the same hypercolumn, units exhibited excitation for very similarly tuned units in the same hypercolumn, but exerted a suppressive effect on units of dissimilar tuning properties, as shown in Fig. 5d. These properties of inter- and intra-columnar connections are roughly consistent with the cooperation between neurons of similar disparities across space (the so-called continuity constraint), and the competition among neurons of different disparities at the same spatial location (the so-called uniqueness constraint) in Marr and Poggio (1976)'s classical stereopsis model for solving the correspondence problem.

However, the lateral connectivity exhibited by the Boltzmann machine model was richer than that in Marr and Poggio (1976)'s model. First, in terms of intra-columnar connections, in addition to the negative (competitive) intra-columnar connections in Marr and Poggio (1976)'s model (Fig. 7a, blue), our model also learned positive intra-columnar connections among units of similar tunings (Fig. 7b). In this aspect, our model is more consistent with

the model in Samonds et al. (2013), which assumes that the intra-columnar interaction has a center excitatory (cooperation between similar neurons) surround inhibitory (competition between dissimilar neurons) profile. This profile is more biologically realistic than that of Marr and Poggio (1976), taking into account the overlapping nature of tuning curves within a hypercolumn, and the model in Samonds et al. (2013) has been shown to explain well the temporal evolution of a number of tuning properties of V1 disparity-tuned neurons.

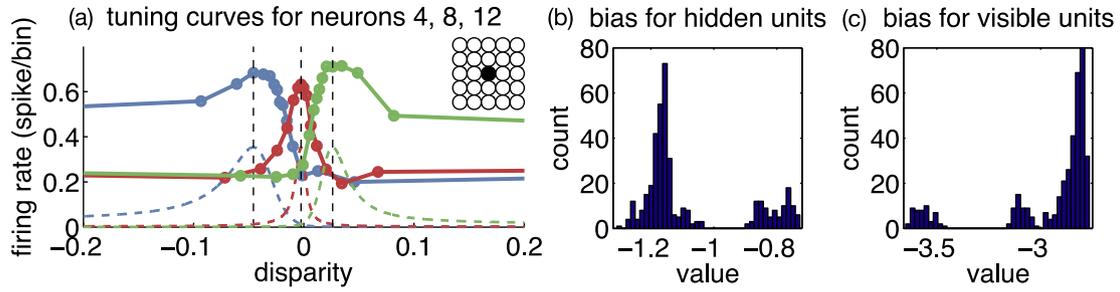
Second, in terms of inter-columnar connections, Marr and Poggio (1976)'s model only specifies positive inter-columnar connections between neurons of the same tuning (Fig. 7a, red), implicitly making the strong assumption that the surfaces of the world are all fronto-parallel. However, surfaces in natural scenes are more diverse, characterized with a variety of surfaces such as slants and tilts, convexities and concavities. This richness in natural scene surface structures likely induced the greater variety of inter-columnar connectivity observed in our model (Fig. 7c) that captures the 3D surface priors to a higher degree than connectivity constraints made in the works of Marr and Poggio (1976) and Samonds et al., 2013. Our model is likely more consistent with more advanced computational models for stereopsis that take into account slant, tilt, and curvature (Li & Zucker, 2010; Belhumeur & Mumford, 1992; Prazdny, 1985).

The learned disparity association fields obviously depend on training natural scene data. Fig. 6 shows the association field obtained by training with data from forest scenes in the Brown data set, with all other parameters unchanged. While the association field along the horizontal direction was symmetrical (Fig. 6a), the one along the vertical direction (Fig. 6b) was skewed, which was not the case for the model trained with interior scenes discussed above (data not shown). This was because the lower parts of forest images are nearer to the viewer than the upper parts, due to the receding ground plane toward the horizon in the forest depth images. This asymmetry along different directions is an interesting prediction that can be tested experimentally.

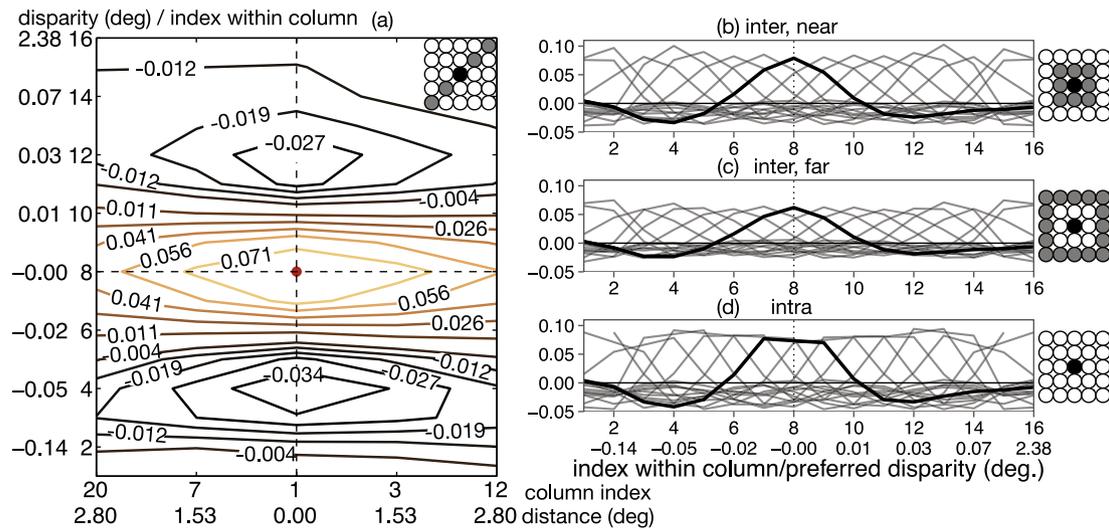
### 3.3. Comparison with neurophysiological data in terms of functional connectivity

We also compared our model with neurophysiological data in terms of functional connectivities derived from CCH measures (Sections 2.2.2 and 2.3.2), and they match qualitatively in three aspects.

First, neurophysiological data in our earlier studies (Samonds et al., 2009) suggested that functional connectivity between a pair of disparity-tuned neurons varied as a function of tuning similarity and the distance between the neuronal pair. Functional connectivity is often measured in terms of the peak of the cross-correlogram (CCH peak) and alternatively the area under the CCH peak within a certain (e.g.  $\pm 10$  ms) window (CCH measure). CCH peak tends to reflect mono-synaptic connections, and CCH measure tends to reflect effective connectivity between a pair of neurons via many possible direct monosynaptic and indirect polysynaptic horizontal or even recurrent feedback connections. Samonds et al. (2009) found that CCH peak and CCH measure were both positively correlated with tuning similarity (measured as the Pearson correlation between the tuning curves), with CCH measure data shown in Fig. 8b. On the other hand, the CCH peak was found to drop with the distance between the centers of the receptive fields, i.e. negatively correlated with RF distance within a visual angle of  $1.5^\circ$ . However, the CCH measure remained relatively constant, i.e. uncorrelated with RF distance within  $1.5^\circ$  (Fig. 8d). This suggested that while the monosynaptic connections between neurons might be fairly local, the effective connections between disparity-tuned neurons are relatively extensive. Last, while functional connectivi-



**Fig. 4.** Some first order properties of the trained Boltzmann machine model. (a) Tuning curves of hidden units 4, 8, 12 in column 1 of the model (inset on top right), shown as solid curves, with corresponding tuning curves of their input units shown as dashed curves. Tuning curves for hidden units were computed using the mean firing rates during simulation at different testing stimuli (dots on curves). (b) Histogram of bias terms for hidden units. (c) Histogram of bias terms for visible units. The cluster of values around  $-0.8$  for hidden units and that around  $-3.6$  for visible units mostly came from units tuned to two ends of the disparity distribution (large negative/positive disparities), which was due to a border effect of the theory in [Ganguli and Simoncelli \(2010\)](#).



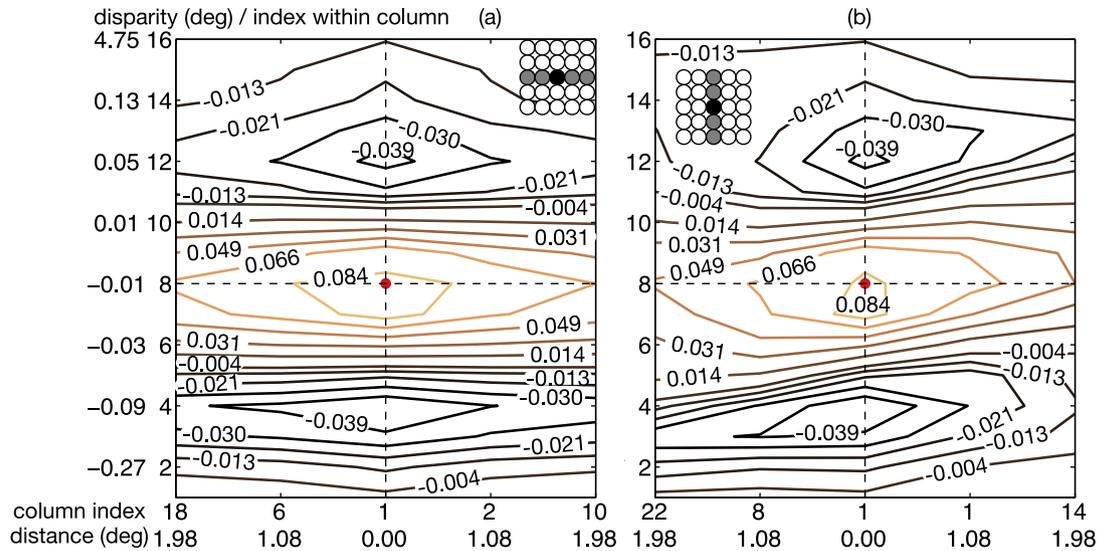
**Fig. 5.** Disparity association field. (a) Disparity association field from unit 8 in the central hypercolumn (column 1) to all other units in 5 hypercolumns (inset on top right; black for column 1, and gray for all 5 but column 1), shown in a contour map of lateral connections. The shapes of contour lines resemble those in the contour association field ([Field et al., 1993](#)). The horizontal axis has two rows, first column index, and second distance to column 1; the vertical axis has two columns, from left to right being the preferred disparity of units, and index of the unit in its hypercolumn (1 to 16). (b) Inter-columnar connections from column 1 to 8 nearby columns. Each curve shows the average connections between one unit in column 1 to unit of a certain index in other 8 columns (inset on right; black for column 1, and gray for other 8). The curve for unit 8 in the central column is highlighted for clarity. Its value at, say, index = 9 is the average of connections from unit 8 in the central column to every unit 9 in surrounding 8 columns, and so on. (c) Inter-columnar connections from column 1 to 16 columns further away (inset on right). The connections were generally weaker than those in the panel above, due to longer distances between columns. (d) Intra-columnar connections within column 1 (inset on right), with the curve for unit 8 highlighted for clarity. Units of very similar tunings tended to facilitate and help each other, but units of different tunings would inhibit each other.

ties existed between neurons of a variety of tuning similarities at close proximity, significant CCH measures could be observed mostly between similarly tuned neurons with RF distance  $> 1^\circ$ , shown in [Fig. 8f](#).

For comparison, [Fig. 8a,c,e](#) show the corresponding results from the Boltzmann machine model. Since there was no synaptic delay in our simulation method, the correlation measures (labeled as CCH measure in [Fig. 8a,c,e](#) for consistency) we obtained from the model were more comparable to the CCH measure, rather than the temporally precise CCH peak in the physiological data. Indeed, we found that the CCH measure positively correlated with tuning similarity ([Fig. 8a](#)) but did not change with RF distance within  $1.5^\circ$  though it did drop gradually over a larger distance ([Fig. 8c](#), pairs with positive CCH measures). This is consistent with the lack of drop in the CCH measure with RF distance in the neural data within  $1.5^\circ$  as shown in [Fig. 8d](#). The model predicted a drop in the CCH measure over a larger RF distance but currently data on long-range CCH measures between disparity-tuned neurons are

not available. However, similar studies by [Smith and Kohn \(2008\)](#) on orientation-tuned neurons did show that the CCH measure dropped only beyond 2 mm in cortical distance between neurons while the CCH peak dropped even at 0.5 mm. Thus, we anticipate that the CCH measure between disparity-tuned neurons will drop beyond  $1.5^\circ$  as our model predicted, but this remains to be tested experimentally. For RF distance vs. tuning similarity ([Fig. 8e](#)), units of similar tuning properties exhibited stronger positive functional connectivities than units of dissimilar tunings when the distance between units became larger.

While all these are in general agreement with the neurophysiological data, there were some differences. Most notably, the model exhibited negative functional connectivities between pairs of units of all distances and similarities ([Fig. 8a,c](#)), whereas the neural data only showed positive connectivities. These differences were largely due to the limitation of the more abstract Boltzmann machine in approximating real neural circuits in both architecture and dynamics as discussed in [Section 4.2](#).



**Fig. 6.** Disparity association fields trained using forest images in the Brown data set. (a) The association field along 5 hypercolumns horizontally. (b) The association field along 5 hypercolumns vertically.

## 4. Discussion

### 4.1. Linking 3D natural scenes and functional connectivity

By training a Boltzmann machine with 3D scene data and simulating neurophysiological experiments on it, this work provides a link between the statistical structure of 3D natural scenes and the measured functional connectivity among disparity-tuned neurons in the primary visual cortex.

We show that certain characteristics of the observed functional connectivity, such as its positive correlation with tuning similarity (Fig. 8b), its decay over receptive field distance between neurons (Fig. 8d), and the prevalence of long range excitatory connections between similarly tuned neurons (Fig. 8f), can be predicted by a Boltzmann machine model trained with natural scene data. The cooperative and competitive connectivity among hidden units in the model is in general agreement with the connectivity constraints in the classical computational model for stereopsis (Marr & Poggio, 1976) and those in the recent neural circuit model of (Samonds et al., 2013).

In previous studies on natural scenes, the learning of cooperation between neurons usually relies on co-occurrence (Field et al., 1993; Geisler et al., 2001; Elder & Goldberg, 2002; Sigman, Cecchi, Gilbert, & Magnasco, 2001; Krüger, 1998), with or without additional supervision signals. Hebbian learning can be used to wire up neurons whose encoded patterns co-occur frequently. Competition is often specified manually or by some hypothetical “anti-Hebbian learning rule”. Our work demonstrates that Boltzmann machines can provide a coherent computational framework to learn facilitatory connections between disparity-tuned neurons based on co-occurrence statistics of disparity signals, as well as the intra-columnar inhibitory connections that implement the uniqueness constraint in earlier computational models (Marr & Poggio, 1976), implemented with “handcrafted” negative connections. Here, we show that if the visual cortex functions mathematically in the manner of a Boltzmann machine or a Markov random field in general, it can acquire these computational constraints—both cooperative and competitive connections—in a unified framework by learning an internal model to explain the input data it experiences during development. The fact that the simulated functional connectivity of the trained Boltzmann machine matches qualitatively with the observed functional connectivity between

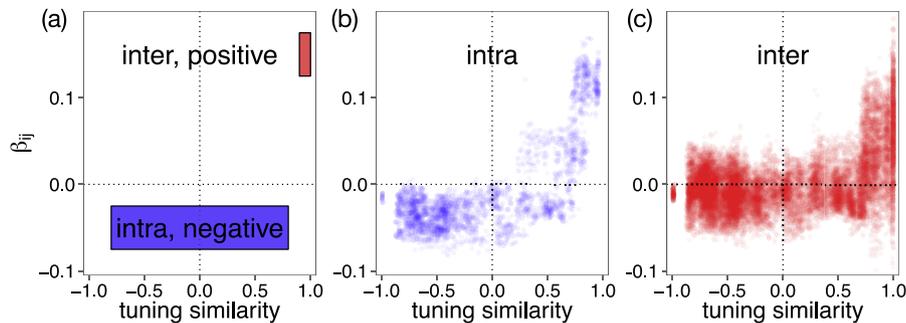
disparity-tuned neurons (Section 3.3) suggests that the interactions among disparity-tuned neurons might form a network, which we call the disparity association field, that encodes the statistics of 3D natural scenes and serves as a prior for solving 3D perception problems. It provides some support to the tantalizing hypothesis that the visual cortex might be functioning like a generative model, e.g. a Boltzmann machine or Markov random field, for statistical inference (Lee & Mumford, 2003; Lee & Yuille, 2006; Lee, 2015).

There is compelling neurophysiological evidence that V1 disparity-tuned neurons are engaged in horizontal recurrent interactions for disambiguation and surface filling-in (Samonds, Tyler, & Lee, 2016). However, the precise biological mechanisms for implementing the Boltzmann machine are not completely clear currently. For the facilitatory connections, Hebbian learning of neurons driven by correlated stimulus signals has been demonstrated in recent *in vivo* rodent experiments (Ko et al., 2011; Ko et al., 2013). Inhibitory connections may be learned by long term depression or homeostatic synaptic scaling mechanisms, and the fairly spatially extensive inhibitory interactions in the model may be implemented through global inhibitory neurons (see Section 4.2 for more discussion).

As an aside and clarification, we want to point out the differences between the functional connectivity (CCH) in neurophysiology (Aertsen et al., 1989; Samonds et al., 2009) and the lateral connection ( $\beta$  in Eq. (4)) in the Boltzmann machine, or other computational models in general (Marr & Poggio, 1976; Samonds et al., 2013). While they are visually comparable in our study (e.g. Figs. 8a vs. 7c), they are mathematically different in nature: CCH reflects correlation whereas  $\beta$  reflects partial correlation or inverse covariance, one (approximately) being the inverse of the other (Friedman, Hastie, & Tibshirani, 2008). The precise relationship between them requires further investigation.

### 4.2. Limitations of the model

One notable discrepancy between the trained Boltzmann machine model and the neural data mentioned in Section 3.3 is that the model had many short-range and long-range negative functional connectivities between hidden units, while the functional connectivity measured between neurons tend to be positive (Fig. 8a,c vs. Fig. 8b,d). There are two possible causes for the discrepancy.



**Fig. 7.** Comparison of connectivity in Marr and Poggio (1976)'s model and in our Boltzmann machine model, in terms of scatter plot of connection vs. tuning similarity between neurons. (a) Schematic of Marr and Poggio (1976)'s model, with negative intra-columnar connections between all neurons of different tunings (blue), and positive inter-columnar connections only between neurons of the same tuning (red). (b) Intra-columnar connections of our model. (c) Inter-columnar connections of our model. The tuning similarity between pairs of hidden units is defined as the Pearson's correlation coefficient between their tuning curves measured from the model simulation (Section 2.2.2) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

First, the input stimuli in the neurophysiological experiments were 8 Hz dynamic random dot stereograms, and the refresh of the stimulus pattern every 125 ms drove the neurons synchronously which could induce a bottom-up positive correlation which might cancel out or overshadow the pairwise negative interaction between neurons.

Second, and possibly more significantly, the brain is not likely to implement extensive local and long-range pairwise negative connections between neurons. Inhibition in the cortical circuitry tends to be mediated by local mechanisms, typically within a hypercolumn. Thus long range inhibition is likely mediated by a cooperation between long-range excitatory connections and local inhibitory neurons. There are numerous types of inhibitory neurons in each hypercolumn, mediating a variety of physiological phenomena such as normalization and surround suppression. The Boltzmann machine offers an interesting proposal, among others, on what mathematical model the cortical connections might be implementing. It would be interesting to explore to what extent the neurons in the visual cortex can implement all the short- and long-range excitations and inhibitions suggested by the Boltzmann machine. In the Boltzmann machine, as in any typical neural network, units can exert both excitation and inhibition on one another. This clearly violates the Dale's law under which a neuron can only be excitatory or inhibitory but not both. There are, however, many examples of recent neural circuit models with excitatory and inhibitory pools of neurons that could be extended to implement neural network models with greater biological realism (King, Zylberberg, & DeWeese, 2013). We are currently investigating an implementation of the model that obeys Dale's law, as well as other realistic biological constraints.

Our Boltzmann machine model was designed to learn only the pairwise connectivity to capture pairwise correlation structures between disparity-tuned neurons across space. It is thus limited in the types of priors that can be encoded. It encodes 3D scene priors in the form of association fields between disparity values across spatial locations, not association fields between 3D surface elements and shapes. Therefore, it cannot, for example, explicitly encode priors on surface structures such as surface slopes (slants and tilts) or surface curvatures. Another layer of hidden units (presumably corresponding to V2 neurons) receiving feedforward connections from disparity-tuned units corresponding to V1 would be required to encode surface priors in the form of the spatial activation patterns in the V1 population. The disparity association field can perform a certain degree of filling-in and surface interpolation as the association field network model showed in Samonds et al. (2013). An association field for completing curved or slant-tilt

surfaces could be implemented in V2 as conjectured in Li and Zucker (2010) and Zucker (2014).

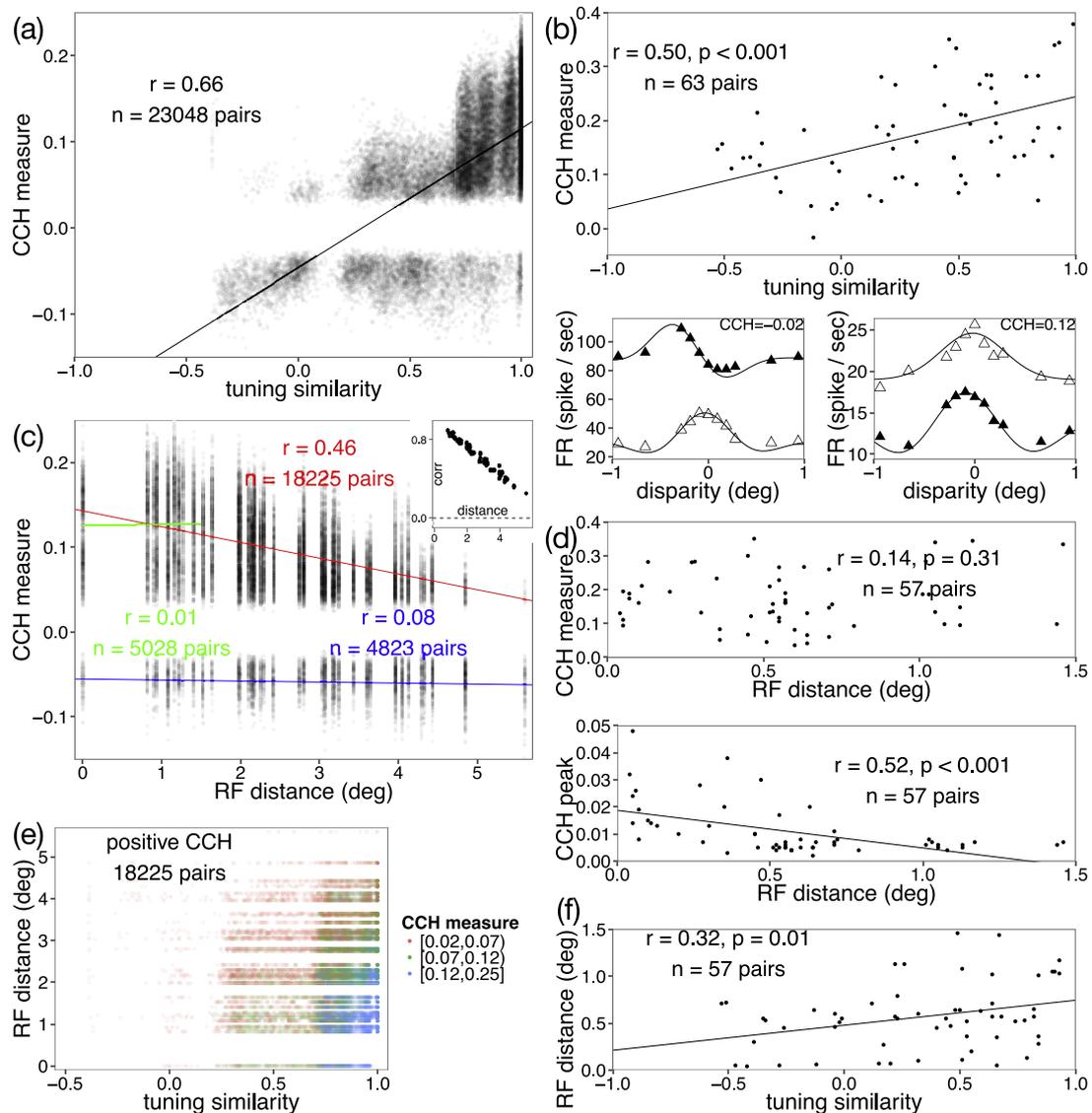
In addition, neurophysiological experiments on our Boltzmann machine model were simulated with Gibbs sampling, which does not strictly follow or exhibit the dynamics of integrate-and-fire neurons in typical neuronal circuit models. However, there have been recent proposals showing that the temporal variability of neuronal spiking activity could potentially be interpreted as Monte Carlo sampling (Hoyer & Hyvärinen, 2003; Fiser, Berkes, Orbán, & Lengyel, 2010). In addition, there have been rigorous mathematical and computational studies on the implementation of sampling algorithms in integrate-and-fire neuronal circuits (Neftci, Das, Pedroni, Kreutz-Delgado, & Cauwenberghs, 2014; Buesing, Bill, Nessler, & Maass, 2011). It is intriguing to contemplate to what extent spike neural networks could actually be implementing the mathematics of the Boltzmann machine.

The comparison between the Boltzmann machine and the neural data we made could only be qualitative for several reasons. We assumed the data set matches the natural experience of the monkeys, and the fixation distribution is uniform. We assumed the distribution of tuning curves for visible units can be derived from the approach in Ganguli and Simoncelli (2010) and that spike trains can be modeled as independent Poisson processes. There are a number of "hyperparameters" in our model, such as the number of units per hypercolumn, and the mean firing rate of units. There is no guarantee that our assumptions are correct, all of which would affect the model's quantitative predictions.

The neural data used were noisy, subject to measurement errors, sampling errors, and are amenable mostly to qualitative comparison. Therefore, we seek mostly to demonstrate that Boltzmann machines can be used to predict qualitatively the functional circuitry of the disparity-tuned neurons in the primary visual cortex, and that natural scenes can predict a general pattern of cooperative and competitive connectivity that we call the "disparity association field". While the learned disparity association field can vary with the assumptions and hyperparameters used quantitatively, the qualitative result in terms of cooperation between similarly tuned units and competition between dissimilarly tuned ones would still hold.

#### 4.3. Summary

The key findings of this paper are as follows. First, certain aspects of cortical circuits can be predicted from Boltzmann



**Fig. 8.** Comparison of model and neurophysiological data in terms of CCH measures. (a)–(b) CCH measure vs. tuning similarity for model (a) and neural data (b). CCH measure and tuning similarity were positively correlated for both cases. The tuning similarity for a pair of neurons (as well as units) is defined as the Pearson's correlation coefficient of their tuning curves, with two examples shown here in the lower part of (b). The right pair of neurons has a larger similarity than the left one, yielding a larger Pearson's correlation. The more similar pair exhibited a stronger CCH measure (0.12 vs. –0.02). (c)–(d) CCH measure vs. distance between neurons for model (c) and neural data (d, upper). In (c), linear fits for pairs with positive (red) and negative (blue) CCH measures are shown separately, and the scatter plot of correlation vs. distance for raw disparity signals is shown in the inset. In addition, a linear fit for positive CCH pairs within  $1.5^\circ$  is shown separately (green) to match the range of distance available in the neural data. The CCH measure of the neural data remained constant within this range of RF distance, i.e. uncorrelated with distance, consistent with the model prediction. In (d), CCH peak vs. distance is also shown in the lower panel. (e)–(f) Distance between neurons vs. tuning similarity for model (e) and neural data (f). For the model, pairs with positive CCH measures were divided into three groups of the same size, and shown in three different colors. Negative CCH pairs were ignored because they were small in magnitude and uniform across different tuning similarities and distances. (b, f) and the lower panel of (d) are adapted from Samonds et al. (2009) with permission of authors.

machines trained on natural scene data. Second, the cortical circuit among disparity-tuned neurons, by virtue of encoding structures in 3D natural scenes, appears to form a disparity association field that could be useful for stereo processing such as removing ambiguity in solving the correspondence problem or performing surface filling-in or interpolation, as some of our recent experiments indicated (Samonds et al., 2016). Third, the structures of the intra-columnar and inter-columnar inhibitory interactions learned by our model suggest that Boltzmann machines might provide an alternative perspective on some prevalent neurophysiological phenomena such as normalization and surround suppression (Carandini & Heeger, 2012); additional studies will be required to confirm this conjecture. Finally, this work suggests

that Boltzmann machines are a viable model for understanding how the Bayesian prior of natural scenes is encoded in the visual cortex. By demonstrating the potential relevance of Boltzmann machines for understanding neural circuitry, this work suggests that a broader class of computational models, called Markov random fields, which are popular and widely used in computer vision and have enjoyed considerable success in solving real early vision problems, might be a viable model of the visual cortex. This work points to the exciting possibility that insights from computer vision on this class of models can be leveraged to understand what problems the visual cortex could be solving and the computational architecture and algorithms underlying the solutions of these problems.

## Acknowledgments

This research is supported by NIH 1R01EY022247, NSF CISE 1320651, NIH NIDA R09 DA023428, NIH P41-EB001977 of the United States, and NSFC 61403090 of China. We gratefully acknowledge the support of NVIDIA Corporation for the donation of GPUs for this research.

## References

- Aertsen, A. M., Gerstein, G. L., Habib, M. K., & Palm, G. (1989). Dynamics of neuronal firing correlation: Modulation of effective connectivity. *Journal of Neurophysiology*, *61*, 900–917.
- Belhumeur, P. N., & Mumford, D. (1992). A Bayesian treatment of the stereo correspondence problem using half-occluded regions. *1992 IEEE computer society conference on computer vision and pattern recognition*. 512: 506.
- Bosking, W. H., Crowley, J. C., & Fitzpatrick, D. (2002). Spatial coding of position and orientation in primary visual cortex. *Nature Neuroscience*, *5*, 874–882.
- Brody, C. D. (1999). Disambiguating different covariation types. *Neural Computation*, *11*, 1527–1535.
- Buesing, L., Bill, J., Nessler, B., & Maass, W. (2011). Neural dynamics as sampling: A model for stochastic computation in recurrent networks of spiking neurons. *PLOS Computational Biology*, *7*, e1002211.
- Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*, 51–62.
- Cohen, M. R., & Maunsell, J. H. R. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*, *12*, 1594–1600.
- Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., Logothetis, N. K., & Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science*, *327*, 584–587.
- Efron, B., Tibshirani, R. J. (1994). *An introduction to the bootstrap*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis.
- Elder, J. H., & Goldberg, R. M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision*, *2*, 324–353.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local association field. *Vision Research*, *33*, 173–193.
- Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: From behavior to neural representations. *Trends in Cognitive Sciences*, *14*, 119–130.
- Fleet, D. J., Wagner, H., & Heeger, D. J. (1996). Neural encoding of binocular disparity: Energy models, position shifts and phase shifts. *Vision Research*, *36*, 1839–1857.
- Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*, 432–441.
- Ganguli, D., & Simoncelli, E. P. (2010). Implicit encoding of prior probabilities in optimal neural populations. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, & A. Culotta (Eds.), *Advances in neural information processing systems* (23, pp. 658–666). Curran Associates Inc.
- Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, *41*, 711–724.
- Geman, S., Geman, D., 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6*, 721–741.
- Gerstein, G. L., & Kirkland, K. L. (2001). Neural assemblies: Technical issues, analysis, and modeling. *Neural Networks*, *14*, 589–598.
- Hinton, G. E., & Sejnowski, T. J. (1986). Learning and relearning in Boltzmann machines. In *Parallel distributed processing: foundations*. MIT Press.
- Hoyer, P. O., & Hyvärinen, A. (2003). Interpreting neural response variability as monte carlo sampling of the posterior. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Advances in neural information processing systems* (15, pp. 300). 293: MIT Press.
- Huang, J., Lee, A. B., & Mumford, D. (2000). Statistics of range images. In *2000 IEEE conference on computer vision and pattern recognition* (pp. 324–331). IEEE Comput. Soc.
- Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron*, *15*, 843–856.
- Kelly, R. C., Smith, M. A., Kass, R. E., & Lee, T. S. (2010). Local field potentials indicate network state and account for neuronal response variability. *Journal of Computational Neuroscience*, *29*, 567–579.
- King, P. D., Zylberberg, J., & DeWeese, M. R. (2013). Inhibitory interneurons decorrelate excitatory cells to drive sparse code formation in a spiking model of V1. *The Journal of Neuroscience*, *33*, 5475–5485.
- Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press.
- Ko, H., Cossell, L., Baragli, C., Antolik, J., Clopath, C., Hofer, S. B., & Mrsic-Flogel, T. D. (2013). The emergence of functional microcircuits in visual cortex. *Nature*, *496*, 96–100.
- Ko, H., Hofer, S. B., Pichler, B., Buchanan, K. A., Sjöström, P. J., & Mrsic-Flogel, T. D. (2011). Functional specificity of local synaptic connections in neocortical networks. *Nature*, *473*, 87–91.
- Koch, C., Marroquin, J., & Yuille, A. (1986). Analog neuronal networks in early vision. *Proceedings of the national academy of sciences*, *83*, 4263–4267.
- Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques*. Adaptive computation and machine learning. MIT Press.
- Krüger, N. (1998). Collinearity and parallelism are statistically significant second-order relations of complex cell responses. *Neural Processing Letters*, *8*, 117–129.
- Lee, T. S. (2015). The visual system's internal model of the world. *Proceedings of the IEEE*, *103*, 1359–1378.
- Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A*, *20*, 1434–1448.
- Lee, T. S., & Yuille, A. L. (2006). Efficient coding of visual scenes by grouping and segmentation. In *Bayesian brain: Probabilistic approaches to neural coding* (pp. 141–185). MIT Press.
- Li, G., & Zucker, S. W. (2010). Differential Geometric Inference in Surface Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*, 72–86.
- Li, W., & Gilbert, C. D. (2002). Global contour saliency and local colinear interactions. *Journal of Neurophysiology*, *88*, 2846–2856.
- Liu, Y., Bovik, A. C., & Cormack, L. K. (2008). Disparity statistics in natural scenes. *Journal of Vision*, *8*, 19–19.
- Marr, D., & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science*, *194*, 283–287.
- Menz, M. D., & Freeman, R. D. (2003). Stereoscopic depth processing in the visual cortex: A coarse-to-fine mechanism. *Nature Neuroscience*, *6*, 59–65.
- Neftci, E., Das, S., Pedroni, B., Kreutz-Delgado, K., & Cauwenberghs, G. (2014). *Event-driven contrastive divergence for spiking neuromorphic systems*. *Frontiers in Neuroscience* (7), .
- Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, *40*, 1392–1405.
- Poggio, G. F., Gonzalez, F., & Krause, F. (1988). Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity. *The Journal of Neuroscience*, *8*, 4531–4550.
- Prazdny, K. (1985). Detection of binocular disparities. *Biological Cybernetics*, *52*, 93–99.
- Samonds, J. M., Potetz, B. R., & Lee, T. S. (2009). Cooperative and competitive interactions facilitate stereo computations in macaque primary visual cortex. *The Journal of Neuroscience*, *29*, 15780–15795.
- Samonds, J. M., Potetz, B. R., & Lee, T. S. (2012). Relative luminance and binocular disparity preferences are correlated in macaque primary visual cortex, matching natural scene statistics. *Proceedings of the National Academy of Sciences*, *109*, 6313–6318.
- Samonds, J. M., Potetz, B. R., Tyler, C. W., & Lee, T. S. (2013). Recurrent connectivity can account for the dynamics of disparity processing in V1. *The Journal of Neuroscience*, *33*, 2934–2946.
- Samonds, J. M., Tyler, C. W., Lee, T. S. (2016). Evidence of stereoscopic surface disambiguation and interpolation in the responses of V1 neurons. Submitted for publication.
- Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences*, *98*, 1935–1940.
- Smith, M. A., & Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. *The Journal of Neuroscience*, *28*, 12591–12603.
- Tappen, M. F., & Freeman, W. T. (2003). Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters. *Proceedings of the Ninth IEEE International Conference on Computer Vision*, 900–906.
- Ventura, V., Cai, C., & Kass, R. E. (2005). Statistical assessment of time-varying dependency between two neurons. *Journal of Neurophysiology*, *94*, 2940–2947.
- Ventura, V., Cai, C., & Kass, R. E. (2005). Trial-to-trial variability and its effect on time-varying dependency between two neurons. *Journal of Neurophysiology*, *94*, 2928–2939.
- von Helmholtz, H. (1896). *Handbuch der physiologischen Optik*. Leipzig Voss.
- Welling, M., & Hinton, G. E. (2002). A new learning algorithm for mean field Boltzmann machines. *International Conference on Artificial Neural Networks*, 2002, 351–357.
- Zucker, S. W. (2014). Stereo, shading, and surfaces: Curvature constraints couple neural computations. *Proceedings of the IEEE*, *102*, 812–829.