Spontaneous emergence of orientation preference and direction selectivity through lateral intracortical interactions

[1]Udo Ernst, [2]Klaus Pawelzik, [3]Misha Tsodyks, and [4]Carmit Sahar-Pikielny

[1]Max-Planck-Institut für Strömungsforschung, D-37073 Göttingen, Germany
[2]Institut für Theoretische Physik, Universität Bremen,
D-28334 Bremen, Germany
[3]Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel
[4]School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv 69978, Israel.

Abstract

The origin and development of receptive field properties of neurons in the primary visual cortex is still under debate. In a simple model consisting of a cortical layer receiving input from oriented stimuli, we show that direction selectivity and orientation preference emerge within milliseconds if lateral interactions within the cortex are strong: activation clusters resulting from a mexican-hat type of interaction are localized through spatial inhomogeneities in the network and in the actual afferent input. Differently oriented stimuli lead to different activation patterns, which are being deformed by the movement of the stimuli. Through this simple mechanism, which provides a new explanation for direction selectivity, maps of direction and orientation preference are formed spontaneously, without any need for a developmental symmetry breaking or a genetic predetermination of the maps of these response properties.

1 Introduction

Neurons in primary visual cortex respond preferentially to oriented gratings or elongated bars moving in one specific direction [11]. Up to now, it is not clear how the neurons achieve their selectivity for these stimulus properties, and what factors determine the layout of spatial maps of orientation and direction preference. A common hypothesis is that neural maps develop through an activity-dependent self-organizing process, and that the layout of these maps is therefore being shaped by the visual experience of the animal. In both experimental and theoretical work, it has been shown that indeed experience can reorganize cortical maps to represent the statistics of a stimulus ensemble [20, 14, 13, 23]. However, recent experiments indicate that neural maps appear very early and independently of visual experience [8, 7]. While one solution for this problem could be a genetic coding of direction and orientation preference, we want to present an alternative hypothesis.
In our approach, we propose a simple model of the visual cortex, consisting of a two-dimensional layer of neuronal units. These units are intracortically coupled with a coupling function having the shape of a mexican hat. Each unit receives input from a specific region of the visual field, where oriented stimuli moving in all possible directions are shown. So far, the model is translationally and rotationally invariant, and there is no direction or orientation tuning coded in the layout of the afferent or lateral weights. Introducing small random inhomogeneities (as e.g. in the lateral weights, afferent weights, neuronal density, etc.) breaks this symmetry, and with the first stimulus ever presented to that network, neurons show up direction and orientation selectivity. Moreover, maps of direction and orientation preference emerge, being a smooth representation with isolated topological defects like pinwheels or fractures, similar to the strutures found in the primary visual cortex [6, 9].
In the next section, we will explain our model in detail. Then, we will show the results and discuss their relevance in the context of recent experimental data.

2 Model

We simulated a two-dimensional rectangular neuronal tissue $\mathcal{C}$ of dimensions L_x by L_y. The density of neurons $\rho(\mathbf{r})$ was assumed to be constant with a small jitter $\eta(\mathbf{r})$, $\rho(\mathbf{r}) = \rho_0 + \eta(\mathbf{r})$, $\eta << \rho_0$. The neuronal activity $f(\mathbf{r})$ evolves according to the inputs received from both geniculate afferents, $I_\mathrm{a}(\mathbf{r})$, and other units in the network, $I_\mathrm{l}(\mathbf{r})$ [22, 10],

$$\frac{\partial f(\mathbf{r})}{\partial t} = -f(\mathbf{r}) + g(I_\mathrm{a}(\mathbf{r}) + I_\mathrm{l}(\mathbf{r})) \quad (1)$$

$$g(I) := s \cdot (I - t_f) \quad \text{for } I > t_f, \text{ and 0 otherwise,} \quad (2)$$

where the piecewise linear gain function g with threshold t_f and slope s models the firing activity in dependence on the total synaptic input current.
Intracortical input is provided by the other units in the network:

$$I_\mathrm{l}(\mathbf{r}) = \int_{\mathcal{C}} \omega(\mathbf{r}, \mathbf{r}') \rho(\mathbf{r}') f(\mathbf{r}') d\mathbf{r}' \quad (3)$$

The interaction strength $\omega(\mathbf{r}, \mathbf{r}')$ between two columns at cortical sites $\mathbf{r}$ and $\mathbf{r}'$ is modeled by a difference of two-dimensional Gaussian functions [4, 17]

$$\omega(\mathbf{r}, \mathbf{r}') = \frac{J_\mathrm{e}}{2\pi {\sigma_\mathrm{e}}^2} \exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|^2}{2{\sigma_\mathrm{e}}^2}\right) - \frac{J_\mathrm{i}}{2\pi {\sigma_\mathrm{i}}^2} \exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|^2}{2{\sigma_\mathrm{i}}^2}\right) \quad . \tag{4}$$

Afferent input to unit j is provided retinotopically by a specific region with center $\mathbf{R}(\mathbf{r}) = \mathbf{r}$ within the visual field $\mathcal{V}$

$$I_{\mathrm{a,j}} = \frac{J_\mathrm{a}}{2\pi {\sigma_\mathrm{a}}^2} \int_{VF} S(\mathbf{R}') \cdot \exp\left(-\frac{\|\mathbf{R}'-\mathbf{R}(\mathbf{r})\|^2}{2{\sigma_\mathrm{a}}^2}\right) d\mathbf{R}' \quad , \tag{5}$$

where $S(\mathbf{R}')$ is the visual stimulus at position $\mathbf{R}'$ in the visual field $\mathcal{V}$. For simplicity, the visual field $\mathcal{V}$ was assumed to have the same dimensions as the cortical surface $\mathcal{C}$. The parameters σ_i, σ_e, and σ_a denote the length scales of the inhibitory and excitatory interaction, and the width of the afferent connection's arbor, respectively. Periodic boundary conditions were used for the cortex and the stimuli to avoid perturbation of the results by geometrical constraints [23].
The network was stimulated by full-sized sinusoidal gratings S_g with period λ, orientation Φ, and velocity v

$$S_g(\mathbf{R}) := S_0 + S_1 \cdot cos\left(\frac{2\pi(x+vt)}{\lambda}\right) \tag{6}$$

$$x := R_\mathrm{x}\cos(\Phi) + R_\mathrm{y}\sin(\Phi) \quad , \tag{7}$$

where S_0 denotes the average luminance, S_1 the contrast of the stimulus, $\mathbf{R} = (R_x, R_y)$ a position in the VF, and t the time.
To obtain direction and angle maps, full-sized gratings moving in $N = 16$ different directions covering the full circle, $\Phi(n) = 2\pi n/N$, were presented to the network. For each of these stimuli, after an initial time interval of $T_0 = 100$ms which was long enough for the activity patterns to build up, one single condition activity map $A_n(\mathbf{r})$, $n = 1, \ldots, N$ has been obtained by averaging the resulting activities $f_n(\mathbf{r}, t)$ over a time interval of $T = 200$ms

$$A_n(\mathbf{r}) = \int_{t=T_0}^{T_0+T} f_n(\mathbf{r}, t)dt \tag{8}$$

Preferred direction $\Theta(\mathbf{r})$ was obtained by vectorially summing up the averaged activities for each direction of motion. Similarly, the preferred orientations $\Phi(\mathbf{r})$ were obtained by the same procedure after first averaging the activity over two opposite directions of motion for each orientation of the grating.

$$\Theta(\mathbf{r}) = \arg\left\{1/N \sum_{n=1}^{N} A_n(\mathbf{r}) \exp(i\Phi(n))\right\} \tag{9}$$

$$\Phi(\mathbf{r}) = \arg\left\{1/N \sum_{n=1}^{N} A_n(\mathbf{r}) \exp(2i\Phi(n))\right\} \tag{10}$$

Our standard parameter set for the simulations shown in the figures was $s = 100\mathrm{Hz}/cdot\mathrm{nA}^{-1}$, $t_f = 0$nA, $J_\mathrm{e} = 45$pAs, $J_\mathrm{i} = 60$pAs, $J_\mathrm{a} = 1$pAs, $S_0 = 1$, $S_1 = 1$,

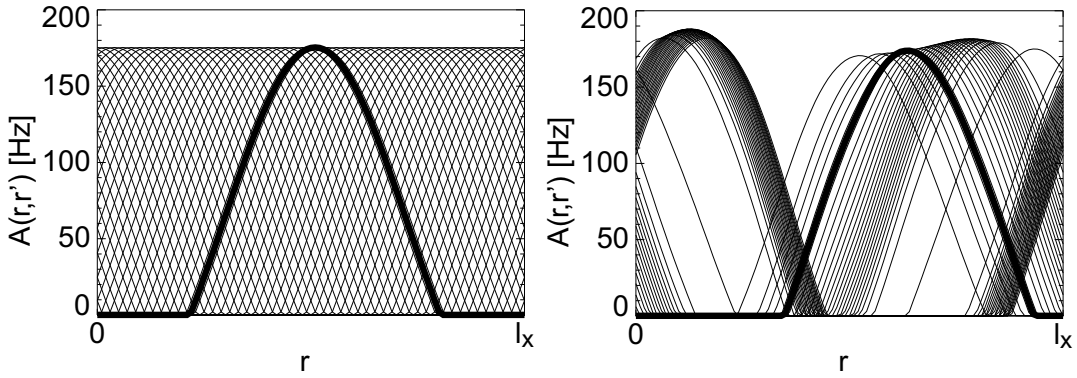

Figure 1: Activation profiles $A(\mathbf{r}, \mathbf{r}')$ of a chain of cortical columns, for input distributions $I_{\mathrm{a}}(\mathbf{r}) = c + \epsilon \exp(-\|\mathbf{r} - \mathbf{r}'\|^2/(2\sigma^2))$, $\epsilon << c$, with $\mathbf{r}'$ equally spaced in the interval $[0, l_x]$. **(a)** With translationally invariant couplings, if the input is peaked around $\mathbf{r}' = l_x/2$, the neuronal response is also centered around $l_x/2$ (thick line). **(b)** With small inhomogeneities in the couplings, the neuronal response is located at a position where afferent **and** lateral input is strongest (thick line). Note that the symmetry breaking leads also to different neuronal response curves, preferring certain cortical regions.

$\lambda = 2^o$, $\sigma_{\mathrm{a}} = 0.25^o$ and $v = 10^o$/s. Basic time step using a fourth-order Runge-Kutta integration scheme was $dt = 0.25$ms. For the numerical simulations, the cortex has been "discretized" into 128 by 128 columns, and the length scales of the intracortical couplings were $\sigma_{\mathrm{e}} = 5.6$ and $\sigma_{\mathrm{i}} = 10$ columns, respectively. The variables of our model are freely scalable, so the units used in this paper are given just to make some connections to biological data.

3 Results

Let us first consider a chain of coupled neuronal units with $\eta(\mathbf{r}) = 0$ and periodic boundary conditions. With a constant input and sufficiently strong intracortical couplings [1], the spatially homogeneous solution of Eq.(1) is unstable. Small inhomogeneities in the input get amplified and lead to the formation of one or more localized cluster(s) of non-zero neuronal activity, the so-called "blobs" [2, 4, 18]. Because of the translation invariance of the network, this solution is marginally stable and the location of the blob is determined by the maxima in the input distribution. If we apply an afferent input which is constant except for a small peak at $\mathbf{r}'$, $I_{\mathrm{a}}(\mathbf{r}) = c + \epsilon \exp(-\|\mathbf{r} - \mathbf{r}'\|^2/(2\sigma^2))$, $\epsilon << c$, then the blob will be peaked around $\mathbf{r}'$ (Fig.(2)(a)).
If $\eta(\mathbf{r}) \neq 0$, some columns will have stronger feedback than others. The symmetry of the system is broken, and the continuum of translation invariant solutions breaks into a discrete set of attractors. If the same afferent input $I_{\mathrm{a}}(\mathbf{r})$ is applied, the position of the blob is now determined both by maximal afferent input **and** by the strength of the cortical feedback [21] (Fig.(2)(b)).
What happens if the input is non-stationary, especially if it is moving with a constant

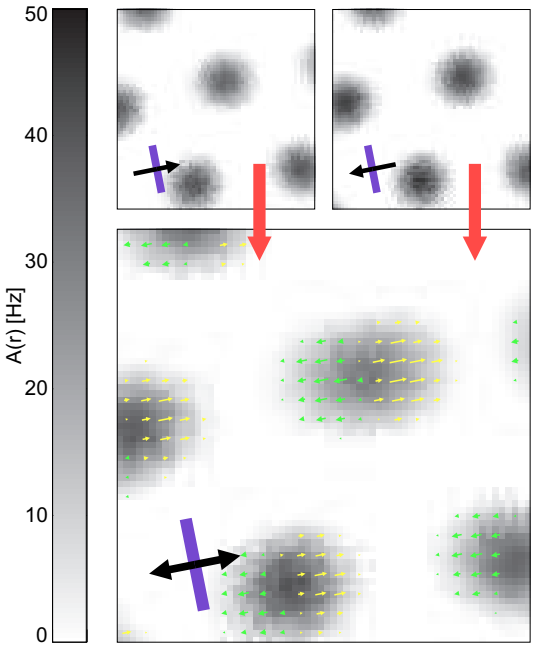

Figure 2: Response $A(\mathbf{r})$ of the network to stimuli moving right (upper left) and to stimuli moving in the opposite direction (upper right). The blobs being at positions of strong cortical feedback are shifted to the right and to the left, respectively. This dynamics leads to a difference in the neuron's response to stimuli moving in opposite directions (below, green and yellow arrows): the columns are direction-selective.

velocity v? Without inhomogeneities (and if v is not too fast), the blob is dragged by the stimulus and moves into the direction of the stimulus movement [5]. With inhomogeneities, the blob is somehow trapped at the position of maximal cortical feedback and moves only slightly. Heuristically, the blob behaves like a particle in a potential well, having not enough energy to climb the walls up to the top. This behaviour can be seen in Fig.(3), where a two-dimensional cortical surface is stimulated with moving gratings S_g. Shifting the stimulus in one of two directions causes the blobs to oscillate around their resting positions. For one column at a fixed position $\mathbf{r}$, this behaviour means that their neuronal response obtained by shifting the stimulus in one direction is less than the response obtained by shifting into the other direction: the unit acquires direction selectivity (Fig.(3)).
While the inhibitory interaction introduces competition leading to the formation of the blobs, the excitatory feedback realizes a form of cooperation among the columns. The coorperation results in the smooth representation of direction preference in a neuronal map (see Fig.(3)). The coupling acts as a filter on the random inhomogeneities, selecting a typical length scale for the direction preference pattern [15, 24], which closely resembles neuronal maps obtained from macaque monkeys. It is obvious that the direction preference map is closely related to the corresponding orientation preference map shown in Fig.(3): each patch of similar orientation preference is divided into two regions of opposite direction preference, as observed in [16].

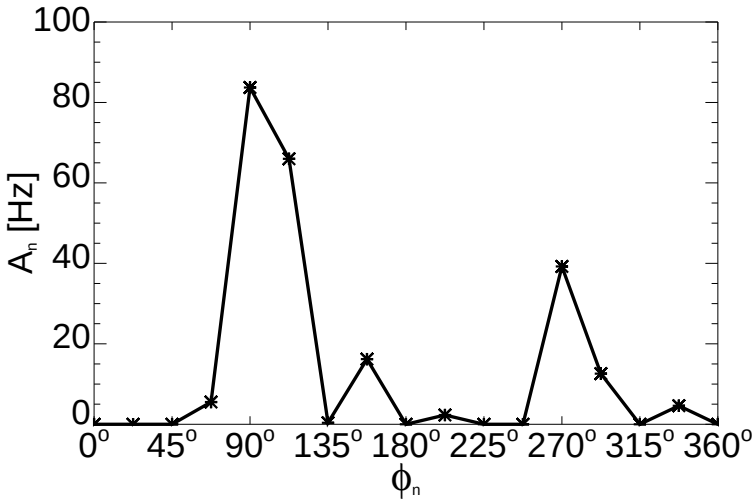

Figure 3: Response A_n of one unit in the model cortex to stimuli of different directions Φ_n. This unit shows an orientation preference of $\Phi = 90^o$ (270^o), and a direction preference of 90^o.

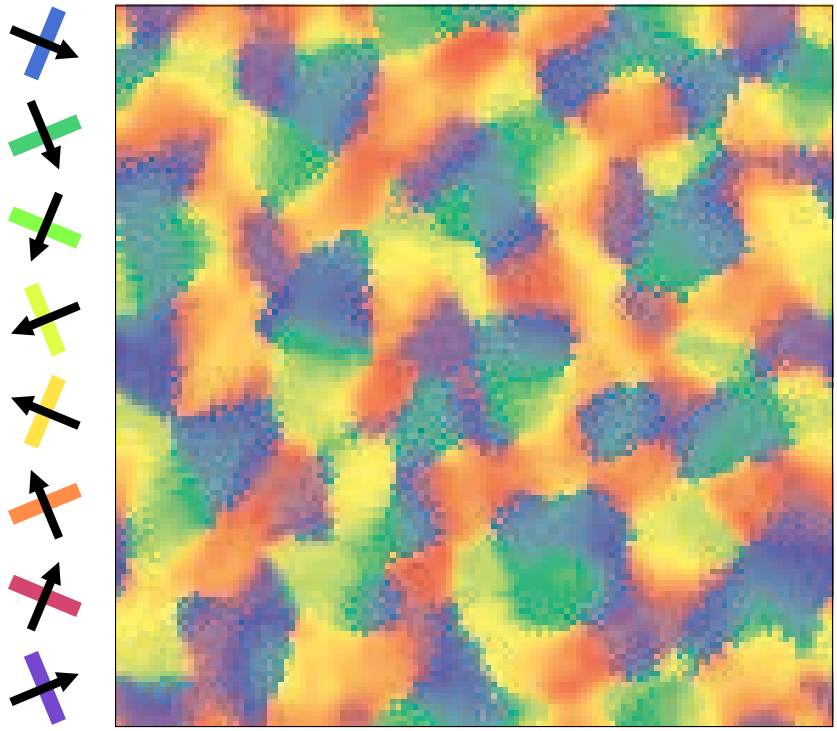

Figure 4: Direction preference map $\Theta(\mathbf{r})$. The preferred direction of each column is coded in the color (see color bars on the left).

4 Discussion

We have seen that it is possible to obtain orientation and direction selectivity in a simple model of the visual cortex. These response properties do not result from

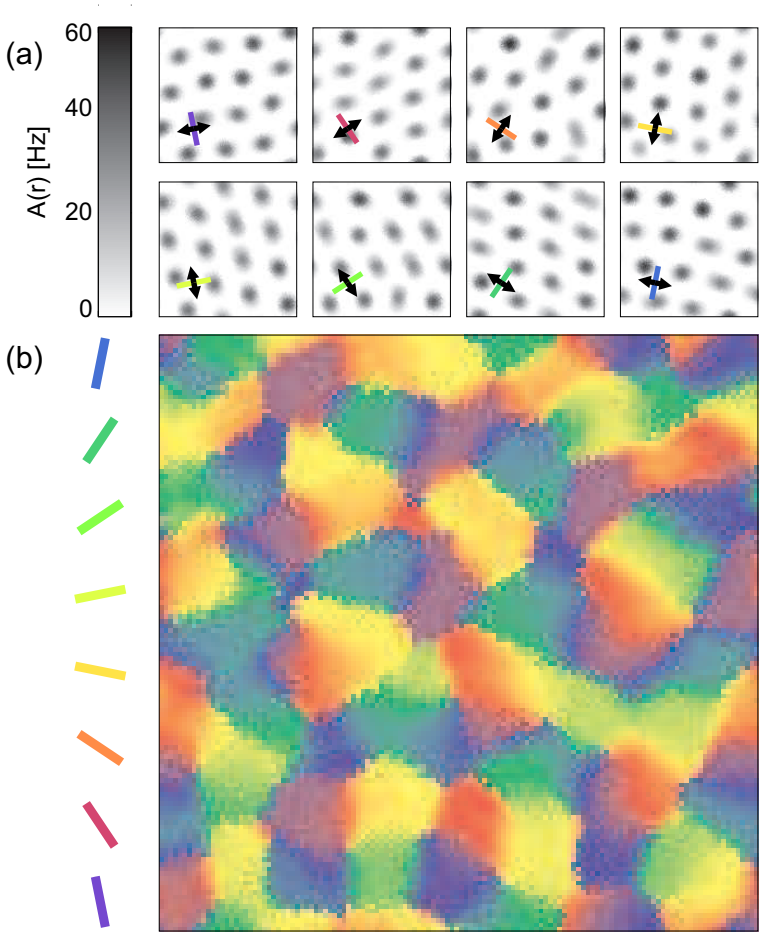

Figure 5: (a) Single condition maps $A_n(\mathbf{r})$ obtained from stimulation with sinusoidal gratings of 8 different orientations. (b) Orientation preference map $\Phi(\mathbf{r})$ obtained by vectorially summing up the single condition maps. The preferred orientation of each column is color coded according to the bars shown on the left. The resulting map is smooth except for small discontinuities like pinwheels or fractures.

a self-organizing developmental process, nor originate from genetic factors hard-wired into the system. The origin of direction preference in our model lies in the small cortical random inhomogeneities inherent to every biological system. With the first stimulus ever presented to that network, these inhomogeneities together with the structure (orientation) of the input get amplified by the intracortical circuitry leading to the observed response properties of the columns.
This framework explains why orientation maps for both eyes are similar, even if the eyes have never had common experience (reverse suture, [8]; binocular deprivation, [7]): Despite the afferent connections from both eyes may be different, the intracortical inhomogeneities determining the blob's positions and the layout of the maps are the same, leading to similar patterns of activation. The instantaneity of the process also provides an explanation for the early presence of orientation preference [11], which develops identically for both eyes for up to three weeks, even if the eyes

were closed throughout this time [7].
Many different mechanisms for the direction selectivity of cortical neurons have been proposed, ranging from a specific arrangement of the synaptic inputs [3], asymmetric dendrites of special cell types [12], up to asymmetries in the afferent inputs amplified by cortical circuitry [19]. Our approach proposes an alternate mechanism relying on **intracortical** random asymmetries which determine the possible positions for activation clusters. Direction selectivity is now dynamically introduced by the movement of the stimulus, which shifts the blobs either in one or the opposite direction. One observation which has already be confirmed by experimental data is that patches of similar orientation preference split into two subpatches of opposite direction selectivity [16]. Another predicition could possibly be tested by similar experiments: the subpatches should be located in the corresponding direction of movement, retinotopically projected onto the retina (see Fig.(3)). We want to emphasize that our approach does not rule out the possibility to reorganize the maps due to visual experience. We speculate that our mechanism arises early, and that the maps are subsequently refined by activity-dependent processes.

Acknowledgments We thank S. Hoshtein, A. Shmuel, S. Löwel, W. Singer and F. Wolf for comments on an earlier draft of the manuscript. This work was supported in part by grants from Max-Planck-Gesellschaft (K.P and M.T), Deutsche Forschungsgemeinschaft (U.E.), Minerva foundation (M.T.), and the Hanse Wissenschaftskolleg (M.T.).

References

[1] B. Ahmed, J. Anderson, R. Douglas, K. Martin, and J. Nelson. Polyneural innervation of spiny stellate neurons in cat visual cortex. *J. Comp. Neurol.*, 341:39–49, 1994.

[2] S. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cyb.*, 27:77–87, 1977.

[3] H. Barlow and W. Levick. The mechanism of directionally selective units in rabbit's retina. *J. Physiol.*, 178:477–504, 1965.

[4] R. Ben-Yishai, R. Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *PNAS*, 92:3844–3848, 1995.

[5] R. Ben-Yishai, D. Hansel, and H. Sompolinsky. Traveling waves and the processing of weakly tuned inputs in a cortical network module. *J. Comp. Neurosci.*, 4:57–77, 1997.

[6] G. Blasdel and G. Salama. Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature*, 321:579–585, 1986.

[7] M. Crair, D. Gillespie, and M. Stryker. The role of visual experience in the development of columns in cat visual cortex. *Science*, 279:566–570, 1998.

[8] I. Gödecke and T. Bonhoeffer. Development of identical maps for two eyes without common visual experience. *Nature*, 379:251–254, 1996.

[9] A. Grinvald, E. Lieke, E. Frostig, R. Frostig, C. Gilbert, and T. Wiesel. Functional architecture of cortex revealed by optical imaging of intrinsic signals. *Nature*, 324:351–354, 1986.

[10] S. Grossberg. Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, 1:17–61, 1988.

[11] D. Hubel and T. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, 160:106–154, 1992.

[12] M. Livingstone. Mechanisms of direction selectivity in macaque $v1$. *Neuron*, 20:509–526, 1998.

[13] K. Miller. A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between on- and off-center inputs of auditory space in the forebrain gaze fields of the barn owl. *J. Neurosci.*, 14:409–441, 1994.

[14] K. Obermayer, H. Ritter, and K. Schulten. A principle for the formation of the spatial structure of cortical feature maps. *PNAS*, 87:8345–8349, 1990.

[15] A. Rojer and E. Schwartz. Cat and monkey cortical columnar patterns modeled by bandpass-filtered 2d white noise. *Biol. Cyb.*, 62:381–391, 1990.

[16] A. Shmuel and A. Grinvald. Functional organization for direction of motion and its relationship to orientation maps in cat area 18. *J. Neurosci.*, 16:6945–6964, 1996.

[17] D. Somers, S. Nelson, and M. Sur. An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.*, 15:5449–5465, 1995.

[18] H. Sompolinsky and R. Shapley. New perspectives on the mechanisms for orientation selectivity. *Cur. Op. Neurobiol.*, 7:514–522, 1997.

[19] H. Suarez, C. Koch, and R. Douglas. Modeling direction selectivity of simple cells in striate visual cortex within the framework of the canonical microcircuit. *J. Neurosci.*, 15:6700–6719, 1995.

[20] N. Swindale, J. Matsubara, and M. Cynader. Surface organization of orientation and direction selectivity in cat area 18. *J. Neurosci.*, 7:1414–1427, 1987.

[21] M. Tsodyks and T. Sejnowski. Associative memory and hippocampal place cells. *Int. Journal of Neural Systems*, 6:81–86, 1995.

[22] H. Wilson and J. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biol. Cyb. (Kybernetik)*, 13:55–80, 1973.

[23] F. Wolf and T. Geisel. Spontaneous pinwheel-annihilation during visual development. *Nature*, 395:73–78, 1998.

[24] F. Wörgötter and E. Niebur. Cortical column desgin: A link between the maps of preferred orientation and orientation tuning strength? *Biol. Cyb.*, 70:1–13, 1993.