Self-Organization and Functional Role of Lateral Connections and Multisize Receptive Fields in the Primary Visual Cortex *

Joseph Sirosh and Risto Miikkulainen
Department of Computer Sciences
University of Texas at Austin
Austin, TX 78712–1188 USA
email: sirosh,risto@cs.utexas.edu

Abstract

Cells in the visual cortex are selective not only to ocular dominance and orientation of the input, but also to its size and spatial frequency. The simulations reported in this paper show how size selectivity could develop through Hebbian self-organization, and how receptive fields of different sizes could organize into columns like those for orientation and ocular dominance. The lateral connections in the network self-organize cooperatively and simultaneously with the receptive field sizes, and produce patterns of lateral connectivity that closely follow the receptive field organization. Together with our previous work on ocular dominance and orientation selectivity, these results suggest that a single Hebbian self-organizing process can give rise to all the major receptive field properties in the visual cortex, and also to structured patterns of lateral interactions, some of which have been verified experimentally and others predicted by the model. The model also suggests a functional role for the self-organized structures: The afferent receptive fields develop a sparse coding of the visual input, and the recurrent lateral interactions eliminate redundancies in cortical activity patterns, allowing the cortex to efficiently process massive amounts of visual information.

1 Introduction

In their first recordings from the primary visual cortex of the cat, Hubel and Wiesel reported that cortical cells were more selective to the width of patterns than were retinal cells [11; 12]. They noted that cortical cells would give no response to a bar covering the whole receptive field (RF), whereas in the retina and the LGN, cells would typically respond to such patterns. Subsequently, detailed studies by Campbell et al. [5], De Valois et al. [7], and others showed that cortical cells are narrowly tuned to the spatial frequency of inputs, and had typical bandpass responses, responding only to inputs in a specific frequency range. A continuum of spatial frequencies from low to high were represented in the cortex [21], and cells in each range of spatial frequency were organized into distinct spatial frequency columns [26; 27]. In essence, cortical cells exhibited an organization of spatial frequency selectivity similar to ocular dominance (OD) and orientation (OR) columns.

Several computational models have been built to demonstrate how other RF properties such as OR preference, OD, and retinotopy can emerge from simple self-organizing processes (e.g. [10; 19;

20; 29]). However, to date, only one computational model has included the development of spatial frequency selectivity. In this so-called Miller's model [18], OR preference and spatial frequency selectivity develop together, and perhaps because of the interactions between these two domains, does not produce a clear columnar organization of spatial frequency selectivity. Although the above models replicate the self-organization of afferent structures quite well, they are based on the simplification that the neuronal response properties are primarily determined by the organization of afferent synapses. Lateral interactions between neurons are approximated by simple mathematical functions (e.g. Gaussians) and assumed to be uniform throughout the network; the structured lateral connectivity of the cortex is not explicitly taken into account. Such models do not explicitly replicate the activity dynamics of the visual cortex, and therefore can make only limited predictions about interactions between receptive fields and cortical function.

Recent experiments have shown that lateral connection patterns closely follow the neuronal response properties [9; 17]. For example, in the normal visual cortex, long-range lateral connections link areas with similar OR preference [9]. Like neuronal response properties, the connectivity pattern is highly plastic in early development and can be altered by experience [13]. Such patterned lateral connections develop at approximately the same time as the cortical columns [4; 13]. Together, these observations suggest that the same experience-dependent process drives the development of both neuronal response properties and lateral connectivity.

Previously, we have shown that a single Hebbian self-organizing process can account for the development of patterned lateral connections, afferent receptive fields, topographic maps and OD columns in the cortex (the Laterally Interconnected Synergetically Self-Organizing Map (LISSOM); [24; 25]). The same algorithm was also shown to explain the organization of orientation maps and the patterns of lateral connections within them [23]. However, we have not studied the selectivity to different-sized stimuli with LISSOM before, although it is a major component of cortical organization. This article investigates whether the same self-organizing process can give rise to RFs selective to different stimulus sizes. Because size selectivity is closely related to spatial frequency selectivity, such self-organization should account for spatial frequency columns as well.

Several new results are reported in this article. It is shown how afferent RFs of different sizes develop from simple retinal images and organize across the network in a systematic fashion. In addition, lateral connections self-organize cooperatively and simultaneously with the size selectivity properties, producing patterns that follow the receptive field organization. In combination with our previous work, these results suggest that a single unified self-organizing process can give rise to not only all the major receptive field properties in the visual cortex, but also the patterns of lateral interactions. The model also suggests a functional role for the lateral interactions: they reduce redundancies in cortical activity and form an efficient sparse coding of the visual input.

2   The Receptive Field LISSOM (RF-LISSOM) model

The LISSOM network is a sheet of interconnected neurons (figure 1). Through afferent connections, each neuron receives input from a "retina". In addition, each neuron has reciprocal excitatory and inhibitory lateral connections with other neurons. Lateral excitatory connections are short-range, connecting only close neighbors. Lateral inhibitory connections run for long distances, and may even implement full connectivity between neurons in the network.

Neurons receive afferent connections from broad overlapping patches on the retina called anatomical RFs. The $N \times N$ network is projected on to the retina of $R \times R$ receptors, and each neuron
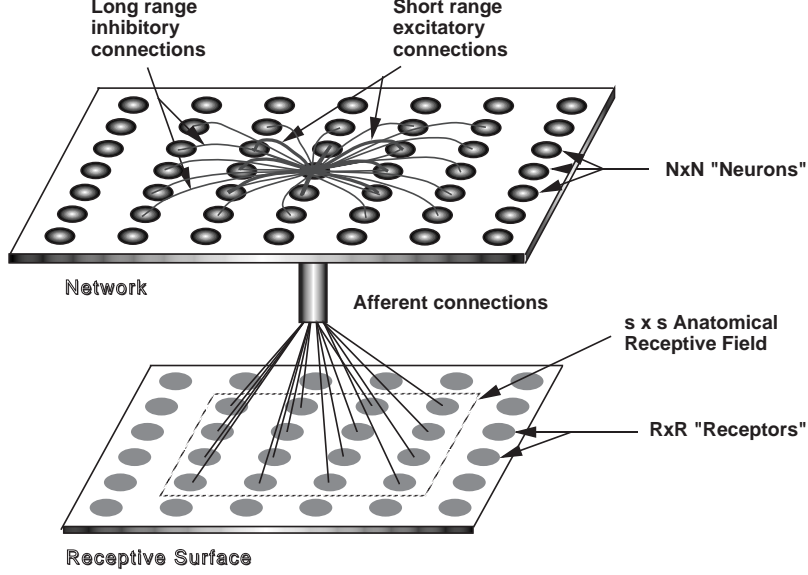
Figure 1: **The Receptive-Field LISSOM architecture.** The afferent and lateral connections of a single neuron in the LISSOM network are shown. The afferents form a local anatomical receptive field on the retina.

is connected to receptors in a square area of side s around the projections. Thus, neurons receive afferents from corresponding regions of the retina. Depending on the location of the projection, the number of afferents to a neuron from the retina could vary from $\frac{1}{2}s \times \frac{1}{2}s$ (at the corners) to $s \times s$ (at the center).

The input to the model consists of gaussian spots of "light" on the retina:

$$\xi_{a,b} = exp\left(-\frac{(a-x_i)^2 + (b-y_i)^2}{u^2}\right) \tag{1}$$

where $\xi_{a,b}$ is the activation of receptor (a,b), u^2 determines the width of the spot, and (x_i, y_i): $0 \leq x_i, y_i < R$ its center. Without normalization, larger-sized spots would produce stronger activation. Therefore, the retinal activity vector is normalized to constant length. The width u is chosen uniformly randomly in a given range, so that inputs of a variety of sizes are presented to the network.

The external and lateral weights are organized through an unsupervised learning process. At each training step, neurons start out with zero activity. The initial response η_{ij} of neuron (i,j) is based on the scalar product

$$\eta_{ij} = \sigma\left(\sum_{a,b} \xi_{ab}\mu_{ij,ab}\right), \tag{2}$$

where ξ_{ab} is the activation of retinal receptor (a,b) within the anatomical RF of the neuron, $\mu_{ij,ab}$ is the corresponding afferent weight, and σ is a piecewise linear approximation of the familiar sigmoid activation function. The response evolves over time through lateral interaction. At each time step, the neuron combines the above afferent activation $\sum \xi\mu$ with lateral excitation and inhibition:

$$\eta_{ij}(t) = \sigma\left(\sum \xi\mu + \gamma_e \sum_{k,l} E_{ij,kl}\eta_{kl}(t-1) - \gamma_i \sum_{k,l} I_{ij,kl}\eta_{kl}(t-1)\right), \tag{3}$$

where $E_{ij,kl}$ is the excitatory lateral connection weight on the connection from neuron (k,l) to neuron (i,j), $I_{ij,kl}$ is the inhibitory connection weight, and $\eta_{kl}(t-1)$ is the activity of neuron
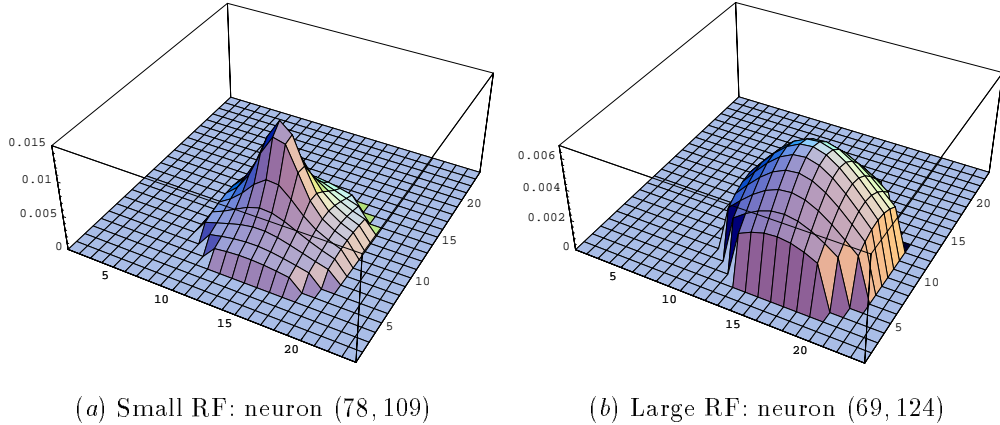
(a) Small RF: neuron $(78, 109)$ (b) Large RF: neuron $(69, 124)$

Figure 2: **Self-organized receptive fields.** The afferent weights of neurons at two different locations in a 192×192 network are shown after self-organization. Initially the weights are completely random, but after self-organization, a smooth hill-shaped weight profile develops. Though the anatomical RFs are the same, the afferent weights are organized into a variety of sizes from narrow, highly peaked receptive fields to large and broad ones.

(k, l) during the previous time step. The constants γ_e and γ_i determine the relative strengths of excitatory and inhibitory lateral interactions. The activity pattern starts out diffuse and spread over a substantial part of the map, and converges iteratively into stable focused patches of activity, or activity bubbles. After the activity has settled, typically in a few iterations of equation 3, the connection weights of each neuron are modified. Both afferent and lateral weights adapt according to the same mechanism: the Hebb rule, normalized so that the sum of the weights is constant:

$$w_{ij,mn}(t + \delta t) = \frac{w_{ij,mn}(t) + \alpha \eta_{ij} X_{mn}}{\sum_{mn} \left[w_{ij,mn}(t) + \alpha \eta_{ij} X_{mn} \right]}, \qquad (4)$$

where η_{ij} stands for the activity of neuron (i, j) in the final activity bubble, $w_{ij,mn}$ is the afferent or lateral connection weight (μ, E or I), α is the learning rate for each type of connection (α_a for afferent weights, α_E for excitatory, and α_I for inhibitory) and X_{mn} is the presynaptic activity (ξ for afferent, η for lateral).

Both inhibitory and excitatory lateral connections follow the same Hebbian learning process and strengthen by correlated activity. At long-distances, very few neurons have correlated activity and therefore most long-range connections eventually become weak. Such weak connections are eliminated periodically, and through weight normalization, inhibition concentrates in a closer neighborhood of each neuron. The radius of the lateral excitatory interactions starts out large, but as self-organization progresses, it is decreased until it covers only the nearest neighbors (c.f. Self-Organizing Map; [14; 15]). Such pruning of lateral connections produces activity bubbles that are focused and local. As a result, weights change in smaller neighborhoods, and receptive fields become better tuned to local areas of the retina.

3   Self-Organization

Simulations were carried out on a network of 192×192 neurons, with inputs coming from a 24×24 retina. The anatomical RF size was chosen to be 11×11, and all the connections were initialized to random weights. A total of $25,000$ training steps were used. At each step, a random-size Gaussian

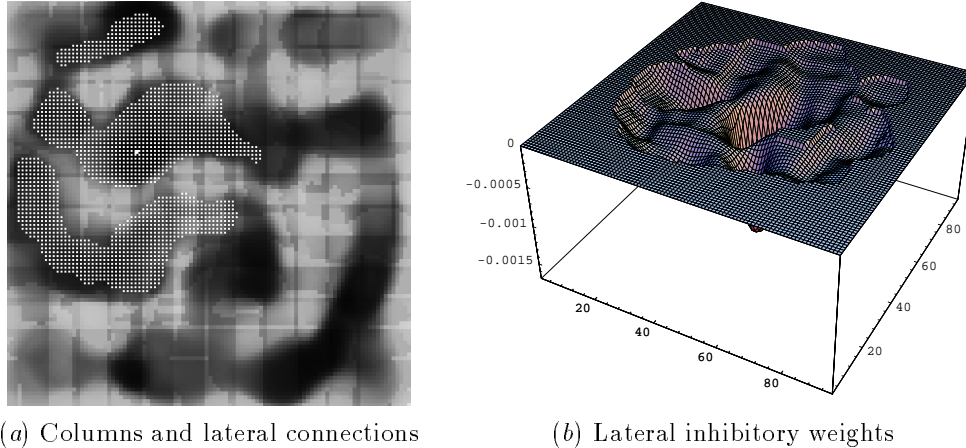(*a*) Columns and lateral connections   (*b*) Lateral inhibitory weights

Figure 3: **Size selective columns and lateral connection patterns.** In figure (*a*), each neuron in the network is labeled with a grey-scale value (*black* → *white*) that represents continuously-changing size preference from small values to large values. Small white dots indicate the lateral input connections to the neuron marked with the big white dot. The size preferences are organized systematically across the network into connected, intertwined patches, and the strongest lateral connections predominantly link areas of the same size selectivity. Figure (*b*) shows the weights of the lateral connections plotted in (*a*). The connection strengths represent the activity correlations of the neuron with the other neurons in the network. The columnar organization of the RFs is reflected in the weights. The connections also are strongest in the immediate vicinity of the neuron (at center) and become weaker with distance. The large areas of zero weights stand for the connections that have been pruned away during self-organization.

spot was presented on the retina as input. The lateral excitatory radius of each neuron started out as 19, but as training progressed, it was gradually decreased to 1. The lateral inhibitory connections had a radius of 47, and weak connections were pruned at intervals of 10,000 iterations.

The self-organization of afferents results in smooth, hill-shaped RFs. A variety of RFs of different sizes are produced, some narrow and tuned to small stimuli, others large and most responsive to large stimuli (figure 2). Simultaneously with the RFs, each neuron's lateral connections evolve, and by the Hebbian mechanism, are distributed according to how well the neuron's activity correlates with the activities of the other neurons (figure 3). Because neurons tuned to similar sizes are frequently active at the same time, the resulting connection strengths are strongest between such neurons and weakest to neurons with very different preferences. The global organization of size preferences and lateral connections can be visualized by labeling each neuron with a color that indicates the width of its RF, and plotting the patterns of lateral connections on top. As figure 3 shows, the RF organization has the form of connected, intertwined patches, similar to OD columns (see e.g. [24]), and the lateral connections of neurons connect to regions of the same size preference.

The columnar organization does not develop in small networks. Simulations show that, for a given variance of the stimuli size, the ratio of neurons in the network to receptors in the retina (the magnification factor) has to be greater than a threshold value for a stable columnar organization to appear. Below the threshold, smooth RFs and an ordered topographic map develop, but all the RFs tend to have the same size, corresponding to the average width of the input stimulus. Above the threshold, symmetry breaking occurs, producing a variety of RF sizes. Such symmetry breaking is similar to that of the Self-Organizing Map [14; 15], where an input feature is represented in the network only if its variance is greater than a threshold proportional to the magnification factor

5

[20].

It is not known whether the long-range lateral connections in the cortex are organized according to size or spatial frequency selectivity. So far, the lateral connection patterns have only been studied in relation to the organization of OD and OR preference [9; 16; 17]. However, considerable psychophysical and neurobiological evidence indicates selective lateral interactions between neurons tuned to different spatial frequencies [3; 6]. As in the RF-LISSOM model, these interactions are also known to be largely inhibitory [6; 28]. The model suggests that the long-range lateral connections could be the anatomical substrate for inhibition between spatial frequency channels. The model further predicts that the patterns of lateral connections in the cortex would be influenced not only by OD and OR preference, but also by selectivity to spatial frequency.

4   Functional role of the self-organized lateral connections and RFs

Combined with our previous work on OD and OR maps and lateral connections, the new results suggest that a single Hebbian mechanism produces the receptive fields and lateral interactions in the primary visual cortex. If so, what could be the functional role of these self-organized structures in visual processing?

Through Hebbian self-organization, the lateral connections learn correlations between the feature detectors in the network—the stronger the correlation between two cells, the larger the connection strength between them. However, these long-range connections are inhibitory. Therefore, the strongly correlated regions of the network inhibit each other—in other words, the lateral connections *decorrelate* [1; 2].

Decorrelation is useful in producing efficient representations. If the connection between two cells is strong, then the response of one can be predicted to a large extent by knowing the response of the other. Therefore, the activity of the second cell is redundant, and a more efficient representation (in an information-theoretic sense) can be formed by eliminating the redundancy. Decorrelation *filters out the learned redundancies* and produces an efficient encoding of the visual input. Thus, the visual knowledge that lateral connections learn is used to filter out the already-known correlations between cortical cells, leaving only novel information to be passed on to higher levels of processing. Our neural network architecture demonstrates how decorrelation mechanisms could be implemented in the primary visual cortex.

The information processing role of the afferent RFs is best seen by analogy with Self-Organizing Maps [14]. The afferent connections self-organize in a similar fashion in both models [22]. When presented with high-dimensional inputs, the self-organizing map selects the set of feature dimensions along which inputs vary the most and represents them along the dimensions of the map [15]. For example, if the inputs lie mostly along the diagonal plane of a hypercube, the self-organized map (and hence the RFs) will spread out along this diagonal. If there is some input variance in the dimension perpendicular to this diagonal, receptive fields will be distributed along this direction as well, and the map will "fold" in that direction. If there are many such feature dimensions, a subset of them will be represented by the folds of the map in the order of their input variance [20]. The images in the visual world could be varying the most along the dimensions of ocular dominance, orientation preference and spatial frequency, and if so, the self-organized RFs will represent these dimensions. During visual processing, the cortex projects incoming visual inputs onto these dimensions. As shown by Field [8], such a projection produces a sparse coding of the

input. Projecting onto the dimensions of maximum variance[1] also achieves minimal distortion and minimal spurious conjunctions of features.

In sum, the RF-LISSOM model predicts that the cortex performs two different computations during sensory processing: First, the inputs are projected onto the principal feature dimensions represented by the afferent receptive field structure. Then, the redundancies are filtered out by recurrent lateral interactions. The result is an efficient, redundancy-reduced sparse coding of the visual input which is then passed on to higher processing levels. This prediction can be verified experimentally by using information theory to analyze the optical images of cortical activity patterns produced in response to simple retinal images. If confirmed, it would constitute a major step in understanding the function of the observed primary visual cortex structures.

5 Conclusion

The RF-LISSOM model shows how a columnar organization of multisize receptive fields can develop and how lateral connection patterns follow this organization. Combined with our previous work, these results show how a single local and unsupervised self-organizing process can be responsible for the development of both the afferent and lateral connection structures in the primary visual cortex. The model suggests that afferent receptive fields develop a sparse coding of the visual input, and that recurrent lateral interactions eliminate redundancies in cortical activity patterns. In essence, the knowledge learned by the lateral connections is used as a negative filter that allows the cortex to efficiently process the massive amounts of visual information presented by the environment.

Acknowledgments

References

[1] H.B. Barlow. Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, 1:371–394, 1972.

[2] H.B. Barlow. Unsupervised learning. *Neural Computation*, 1:295–311, 1989.

[3] L.A. Bauman and A.B. Bonds. Inhibitory refinement of spatial frequency selectivity in single cells of the cat striate cortex. *Vision Research*, 31(6):933–944, 1991.

[4] A. Burkhalter, K. L. Bernardo, and V. Charles. Development of local circuits in human visual cortex. *Journal of Neuroscience*, 13:1916–1931, May 1993.

[5] F.W. Campbell, G.F. Cooper, and C. Enroth-Cugell. The spatial selectivity of the visual cells of the cat. *Journal of Physiology (London)*, 203:223–235, 1969.

[6] K.K. De Valois and R.B.H. Tootell. Spatial-frequency-specific inhibition in cat striate cortex cells. *Journal of Physiology (London)*, 336:359–376, 1983.

---

[1]The dimensions of maximum variance are not necessarily those given by linear principal component analysis of the input, as shown by Field [8].

[7] R.L. De Valois, D.G. Albrecht, and L.G. Thorell. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22:545–559, 1982.

[8] D. J. Field. What is the goal of sensory coding? *Neural Computation*, 6:559–601, 1994.

[9] C.D. Gilbert and T.N. Wiesel. Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *Journal of Neuroscience*, 9:2432–2442, 1989.

[10] G. Goodhill. Topography and ocular dominance: a model exploring positive correlations. *Biological Cybernetics*, 69:109–118, 1993.

[11] D.H. Hubel and T.N. Wiesel. Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology*, 148:574–591, 1959.

[12] D.H. Hubel and T.N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology (London)*, 160:106–154, 1962.

[13] L.C. Katz and E.M. Callaway. Development of local circuits in mammalian visual cortex. *Annual Review of Neuroscience*, 15:31–56, 1992.

[14] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69, 1982.

[15] T. Kohonen. *Self-Organization and Associative Memory*. Springer, Berlin; Heidelberg; New York, third edition, 1989.

[16] S. Löwel and W. Singer. Selection of intrinsic horizontal connections in the visual cortex by correlated neuronal activity. *Science*, 255:209–212, 1992.

[17] R. Malach, Y. Amir, M. Harel, and A. Grinvald. Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in the primate striate cortex. *Proceedings of the National Academy of Sciences, USA*, 90:10469–10473, 1993.

[18] K.D. 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. *Journal of Neuroscience*, 14:409–441, 1994.

[19] K.D. Miller, J.B. Keller, and M.P. Stryker. Ocular dominance column development: Analysis and simulation. *Science*, 245:605–615, 1989.

[20] K. Obermayer, G.G. Blasdel, and K.J. Schulten. Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps. *Physical Review A*, 45:7568–7589, 1992.

[21] M.S. Silverman, D.H. Grosof, R.L. De Valois, and S.D. Elfar. Spatial frequency organization in primate striate cortex. *Proceedings of the National Academy of Sciences, USA*, 86:711–715, 1989.

[22] J. Sirosh and R. Miikkulainen. Cooperative self-organization of afferent and lateral connections in cortical maps. *Biological Cybernetics*, 71:66–78, 1994.

[23] J. Sirosh and R. Miikkulainen. Cooperative self-organization of orientation maps and lateral connections in the visual cortex. In *Society for Neuroscience Abstracts*, volume 21, Part 3, page 1751, 1995.

[24] J. Sirosh and R. Miikkulainen. Ocular dominance and patterned lateral connections in a self-organizing model of the primary visual cortex. In G. Tesauro, D.S. Touretzky, and T.K. Leen, editors, *Advances in Neural Information Processing Systems 7*, Cambridge, MA, 1995. MIT Press.

[25] J. Sirosh and R. Miikkulainen. Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex. *Neural Computation*, in press.

[26] R.B. Tootell, M.S. Silverman, and R.L. De Valois. Spatial frequency columns in primary visual cortex. *Science*, 214:813–815, November 13 1981.

[27] R.B. Tootell, M.S. Silverman, S.L. Hamilton, E.Switkes, and R.L. De Valois. Functional anatomy of macaque striate cortex. V. Spatial frequency. *Journal of Neuroscience*, 8:1610–1624, May 1988.

[28] T.R. Vidyasagar and A. Mueller. Function of GABA inhibition in specifying spatial frequency and orientation selectivities in cat striate cortex. *Experimental Brain Research*, 98(1):31–38, 1994.

[29] C. von der Malsburg. Self-organization of orientation-sensitive cells in the striate cortex. *Kybernetik*, 15:85–100, 1973.