Topographic Receptive Fields and Patterned Lateral Interaction in a Self-Organizing Model of The Primary Visual Cortex

Joseph Sirosh and Risto Miikkulainen
Department of Computer Sciences
The University of Texas at Austin, Austin, TX 78712
{sirosh,risto}@cs.utexas.edu

Abstract

A self-organizing neural network model for the simultaneous and cooperative development of topographic receptive fields and lateral interactions in cortical maps is presented. Both afferent and lateral connections adapt by the same Hebbian mechanism in a purely local and unsupervised learning process. Afferent input weights of each neuron self-organize into hill-shaped profiles, receptive fields organize topographically across the network, and unique lateral interaction profiles develop for each neuron. The model demonstrates how patterned lateral connections develop based on correlated activity, and explains why lateral connection patterns closely follow receptive field properties such as ocular dominance.

1 Introduction

The response properties of neurons in many sensory cortical areas are ordered topographically, that is, nearby neurons respond to nearby areas of the receptor surface. Such topographic maps form by the self-organization of afferent connections to the cortex, driven by external input (Hubel and Wiesel 1965; Miller et al. 1989; Stryker et al. 1988; von der Malsburg 1973). Several neural network models have demonstrated how global topographic order can emerge from local cooperative and competitive lateral interactions within the cortex (Amari 1980; Kohonen 1982, 1993; Miikkulainen 1991; Willshaw and von der Malsburg 1976). These models are based on predetermined lateral interaction and focus on explaining how the afferent synaptic weights are organized.

A number of recent neurobiological experiments indicate that lateral connections self-organize like the afferent connections: (1) The lateral connectivity is not uniform or genetically predetermined, but forms during the early development based on external input (Katz and Callaway 1992; Löwel and Singer 1992). (2) In the primary visual cortex, lateral connections are initially widespread, but develop into clustered patches. The clustering period overlaps substantially with the period during which ocular dominance and orientation columns form (Katz and Callaway 1992; Dalva and Katz 1994; Burkhalter et al. 1993). (3) Lateral connections primarily connect areas with similar response properties, such as columns with the same orientation or (in the strabismic case) eye preference (Gilbert 1992; Löwel and Singer 1992). (4) The lateral connections are far more numerous than the afferents and are believed to have a substantial influence on cortical activity (Gilbert et al. 1990). To fully account for cortical self-organization, a cortical map model must demonstrate that both afferent and lateral connections can organize simultaneously, from the same external input, and in a mutually supportive manner.

To appear in *Neural Computation*.

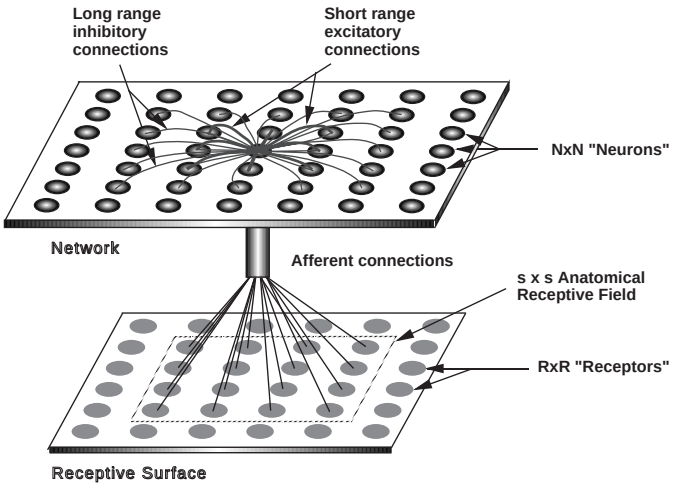


Figure 1: **Architecture of the self-organizing network.** The lateral excitatory and lateral inhibitory connections of a single neuron in the network are shown, together with its afferent connections. The afferents form a local anatomical receptive field on the retina.

We have previously shown how Kohonen's self-organizing feature maps (Kohonen 1982) can be generalized to include self-organizing lateral connections (the Laterally Interconnected Synergetically Self-Organizing Map (LISSOM); Sirosh and Miikkulainen 1993, 1994). LISSOM is a low-dimensional abstraction of the cortical self-organizing process and models a small region of the cortex where all neurons receive the same input vector. In contrast, this paper shows how realistic, high-dimensional receptive fields develop as part of the self-organization, and in essence scales up the LISSOM approach to large areas of the cortex where different parts of the cortical network receive inputs from different parts of the receptor surface. This Receptive-Field LISSOM model shows how (1) topographically ordered receptive fields develop from simple retinal images, (2) lateral connections self-organize cooperatively and simultaneously with the afferents, (3) long-range lateral connections store correlations in activity across the topographic map, and (4) the resulting lateral connection patterns closely follow receptive field properties such as ocular dominance.

2 The Receptive-Field LISSOM Model

The cortical network is modeled as a sheet of neurons interconnected by short-range excitatory lateral connections and long-range inhibitory lateral connections (figure 1). Neurons receive input from a receptive surface or "retina" through the afferent connections. These connections come from overlapping patches on the retina called anatomical receptive fields, or RFs. The patches are distributed with a given degree of randomness. The $N \times N$ network is projected on the retina of $R \times R$ receptors, and each neuron is assigned a receptive field center (c_1, c_2) randomly within a radius $\rho * R$ ($\rho \in [0, 1]$) of the neuron's projection. Through the afferent connections, the neuron receives input from receptors in a square area around the center with side s. Depending on its location, the number of afferents to a neuron could vary from $\frac{1}{2}s \times \frac{1}{2}s$ (at the corners) to $s \times s$ (at the center).

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

$$\eta_{ij} = \sigma\left(\sum_{r_1,r_2} \xi_{r_1,r_2}\mu_{ij,r_1r_2}\right), \tag{1}$$

where ξ_{r_1,r_2} is the activation of a retinal receptor (r_1, r_2) within the receptive field of the neuron, μ_{ij,r_1r_2} is the corresponding afferent weight, and σ is a piecewise linear approximation of the familiar sigmoid activation function:

$$\sigma(x) = \begin{cases} 0 & x \leq \delta \\ (x-\delta)/(\beta-\delta) & \delta < x < \beta \\ 1 & x \geq \beta \end{cases} \tag{2}$$

where δ and β are the lower and upper thresholds. The response evolves over time through lateral interactions. At each time step, the neuron combines retinal activation with lateral excitation and inhibition:

$$\eta_{ij}(t) = \sigma\left(\sum_{r_1,r_2} \xi_{r_1,r_2}\mu_{ij,r_1r_2} + \gamma_e \sum_{k,l} E_{ij,kl}\eta_{kl}(t-\delta t) - \gamma_i \sum_{k,l} I_{ij,kl}\eta_{kl}(t-\delta t)\right), \tag{3}$$

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 $\eta_{kl}(t-\delta t)$ is the activity of neuron (k,l) during the previous time step. All connection weights are positive. The input activity stays constant while the neural activity settles. The scaling factors γ_e and γ_i determine the strength of the lateral excitatory and inhibitory interactions. The activity pattern starts out diffuse and spread over a substantial part of the map, but within a few iterations of equation 3, converges into a stable focused patch of activity, or activity bubble. After the activity has settled, the connection weights of each neuron are modified. Both afferent and lateral connection weights adapt according to the same mechanism: the Hebb rule, normalized so that the sum of the weights is constant:

$$w_{ij,mn}(t+1) = \frac{w_{ij,mn}(t) + \alpha\eta_{ij}X_{mn}}{\sum_{mn}\left[w_{ij,mn}(t) + \alpha\eta_{ij}X_{mn}\right]}, \tag{4}$$

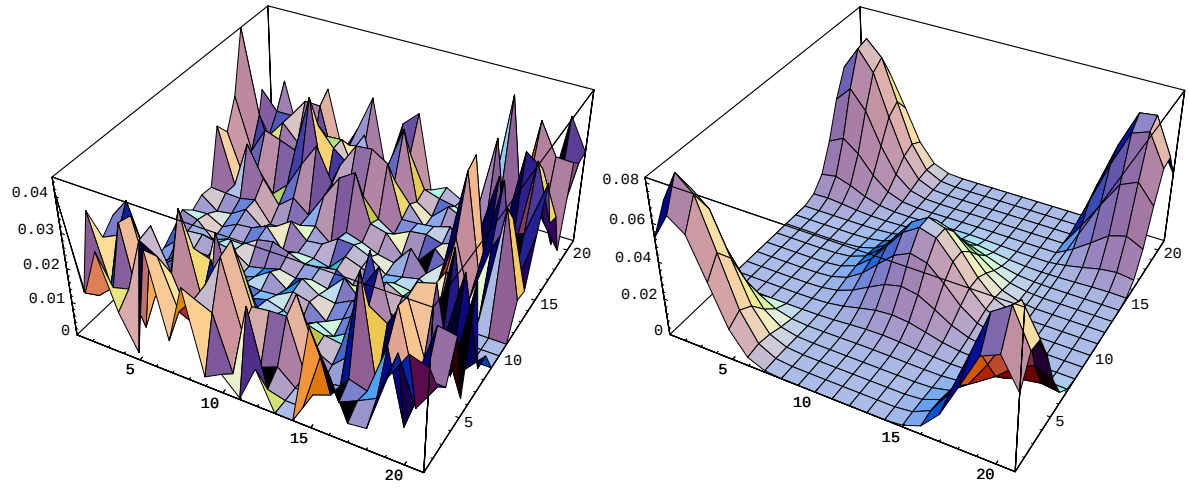
where η_{ij} stands for the activity of the neuron (i,j) in the settled activity bubble, $w_{ij,mn}$ is the afferent or the lateral connection weight (μ_{ij,r_1r_2}, $E_{ij,kl}$ or $I_{ij,kl}$), α is the learning rate for each type of connection (α_A for afferent weights, α_E for excitatory, and α_I for inhibitory) and X_{mn} is the presynaptic activity (ξ_{r_1,r_2} for afferent, η_{kl} for lateral). Afferent inputs, lateral excitatory inputs, and lateral inhibitory inputs are normalized separately. The larger the product of the pre- and post-synaptic activity $\eta_{ij}X_{mn}$, the larger the weight change. Therefore, both excitatory and inhibitory connections strengthen by correlated activity; normalization then redistributes the changes so that the sum of each weight type for each neuron remains constant.

3 Self-organization

Although the self-organizing mechanism outlined above is robust, the width of the anatomical receptive fields and how ordered they are strongly affect the outcome of the process. In a series of simulations, the conditions under which self-organization could take place were studied, using networks with various receptive field widths and varying degrees of initial order. In each case, all synaptic weights were initially random: a uniformly distributed random value between zero and one was assigned to all weights, and the total weight of each connection type of each neuron was normalized to 1.0.[1] Gaussian spots of "light" on the retina were used as input. At each presentation, the activation ξ_{r_1,r_2} of the receptor (r_1, r_2) was given by:

$$\xi_{r_1,r_2} = \sum_{i=1}^{n} exp(-\frac{(r_1-x_i)^2 + (r_2-y_i)^2}{a^2}) \tag{5}$$

[1] Various schemes for initializing afferent connections have been studied by other researchers. These schemes include tagging connections with chemical markers (Willshaw and von der Malsburg 1979) and establishing an initial topographic bias (Willshaw and von der Malsburg 1976; Goodhill 1993). Our focus is on receptive field width vs. initial order because this factor has turned out most crucial in determining the success of the self-organizing process in the present model.



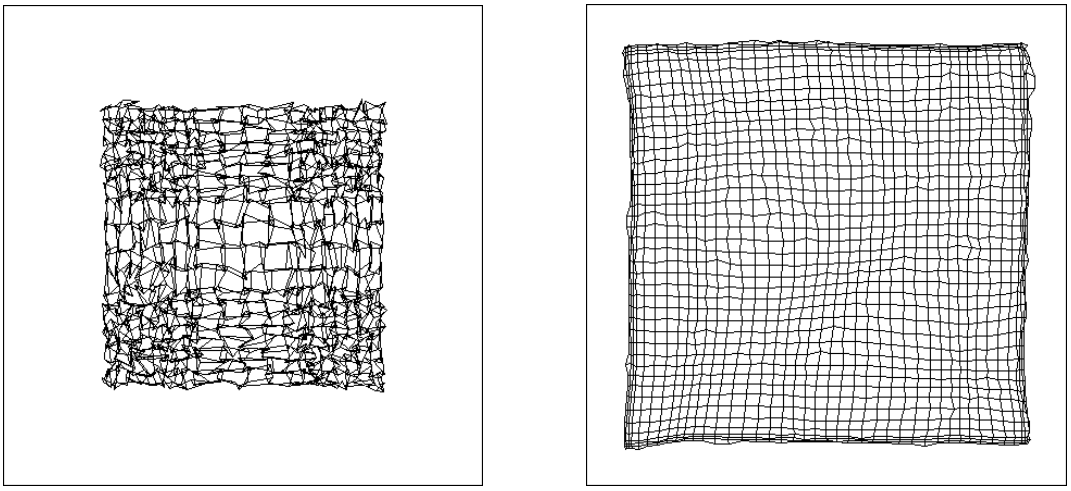
(*a*) Initial random weights (*b*) Final organized receptive fields

Figure 2: **Self-organization of the afferent input weights.** The afferent weights of five neurons (located at the center and at the four corners of the network) are superimposed on the retinal surface in this figure. The retina had 21×21 receptors, and the receptive field radius was chosen to be 8. Therefore, neurons could have anywhere from 8×8 to 17×17 afferents depending on their distance from the network boundary. (*a*) The anatomical RF centers were slightly scattered around their topographically ordered positions (uniformly, within a distance of 0.5 in retinal coordinates), and the weights were initialized randomly (as discussed in the text). There are four concentrated areas of weights slightly displaced from the corners, and one larger one in the middle. At the corners, the profiles are taller because the normalization divides the total afferent weight among a smaller numer of connections. (*b*) As the self-organization progresses, the weights organize into smooth hill-shaped profiles. In this simulation, each input consisted of 3 randomly-located Gaussian spots with $a = 2.0$. The lateral interaction strengths were $\gamma_e = \gamma_i = 0.9$, the learning rates $\alpha_{\rm A} = \alpha_{\rm E} = \alpha_{\rm I} = 0.002$, and the upper and lower thresholds of the sigmoid 0.65 and 0.1. The map was formed in 10,000 training presentations.

where n is the number of spots, a^2 specifies the width of the Gaussian, and the spot centers (x_i,y_i): $0 \leq x_i, y_i < R$, were chosen randomly.

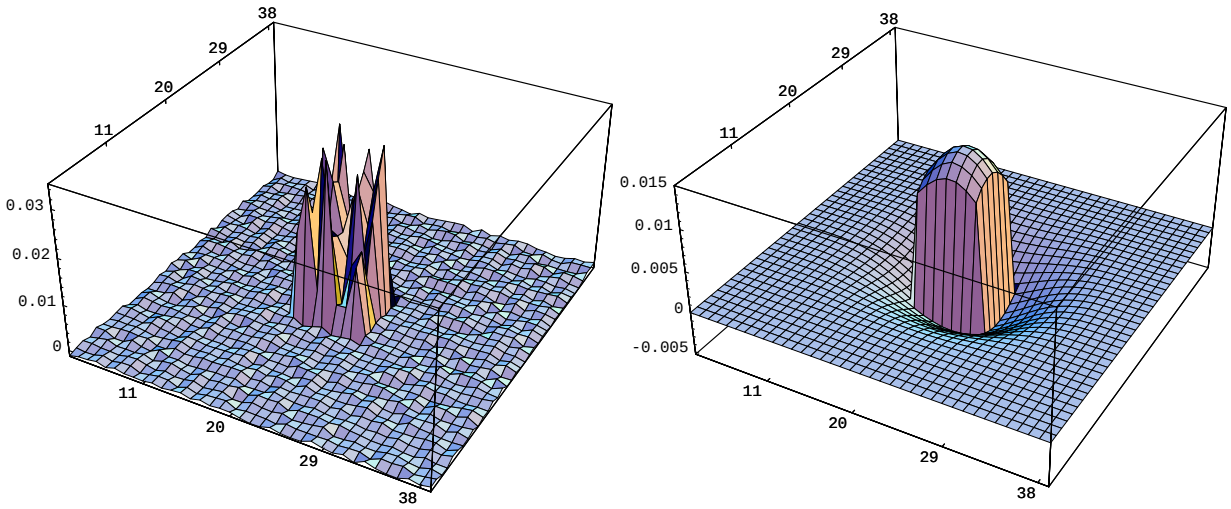
When the networks were trained with single light spots ($n = 1$), similar afferent and lateral connection structures developed in all cases. With more realistic input consisting of multiple light spots ($n > 1$), the networks with wide anatomical receptive fields and relatively small topographic scatter self-organized just as robustly. Figures 2—4 illustrate the self-organization of such a network. The initial rough pattern of afferent weights of each neuron evolved into a hill-shaped profile (figure 2). The afferent weight profiles of different neurons peaked over different parts of the retina, and their center of gravities (calculated in retinal coordinates) formed a topographical map (figure 3).

Networks with large topographic scatter compared to the RF size, however, failed to develop global order with multi-spot inputs. It is interesting to analyze why. With single light spots, the afferent weights of an active neuron always change towards a single, local input pattern. Eventually these weights become concentrated around a local area in the RF such that the global distribution of the centers best approximates the input distribution. However, when multiple spots occur in the receptive field at the same time, the weights change towards several different locations. If the scatter is large and the neighboring receptive fields do not overlap substantially, the inputs in the nonintersecting areas change the weights the most, and the receptive fields remain crossed. If the scatter is small and the overlap high, the input activity in the intersecting region becomes strong enough to drive the weights towards topographic order. The model therefore suggests that for the self-organization of afferent connections in the cortex to be purely activity-driven, the anatomical



(*a*) Initial slightly disordered map (*b*) Final ordered map

Figure 3: **Self-organization of afferent receptive fields into a topographic map.** The center of gravity of the afferent weight vector of each neuron in the 42×42 network is projected onto the receptive surface (represented by the square). Each center of gravity point is connected to those of the four immediately neighboring neurons by a line. The resulting grid depicts the topographical organization of the map. Initially, the anatomical RF centers were only slightly scattered topographically, but because the afferent weights were initially random, the centers of gravity are scattered much more (*a*). As the self-organization progresses, the map unfolds and the weight vectors spread out to form a regular topographic map of the receptive surface, such as shown in (*b*). Because the anatomical receptive fields are large in this example compared to the initial topographic scatter, the self-organization is robust both with single and multi-spot inputs.



(*a*) Initial rough interaction profile (*b*) Final smooth interaction

Figure 4: **Self-organization of the lateral interaction.** The lateral interaction profile for a neuron at position $(20, 20)$ in the 42×42 network is plotted. The excitation and inhibition weights are initially randomly distributed within radii 3 and 18. The combined interaction is the sum of the excitatory and inhibitory weights and illustrates the total effect of the lateral connections. The sums of excitation and inhibition were chosen to be equal, but because there are fewer excitatory connections, the interaction has the shape of a rough plateau with a central peak (*a*). During self-organization, smooth patterns of excitatory and inhibitory weights evolve, resulting in a smooth "Mexican hat" shaped lateral interaction profile (*b*).

receptive fields must be large compared to the amount of initial topographic scatter among their centers.[2]

The lateral connections evolve together with the afferents. By the normalized Hebbian rule (equation 4), the lateral connection weights of each neuron are distributed according to how well the neuron's activity correlates with the activities of the other neurons. As the afferent receptive fields organize into a uniform map (figure 3), these correlations fall off with distance approximately like a Gaussian, with strong correlations to near neighbors and weaker correlations to more distant neurons. The lateral excitatory and inhibitory connections acquire the Gaussian shape, and the combined lateral excitation and inhibition becomes an approximate difference of Gaussians (or a "Mexican hat"; figure 4).

4 Ocular dominance and lateral connection patterns

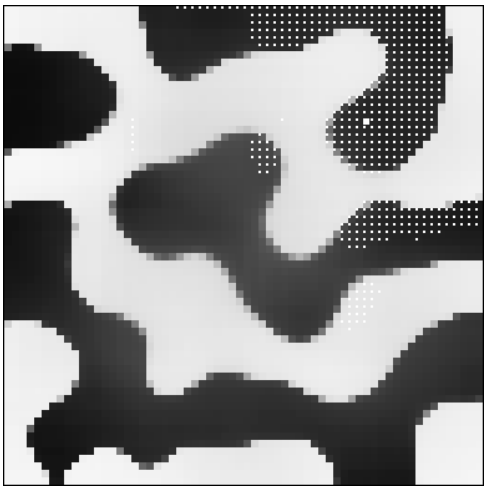
The activity patterns and correlations in the retina are not as simple or random as in the above experiments. Also, the visual cortex has a complex organization of orientation and ocular dominance columns, and this organization is reflected in the patterns of lateral connections. As a first step to studying how afferent and lateral connections evolve in the visual cortex, the development of ocular dominance was simulated. A second retina was added to the model and afferent connections were set up exactly as for the first retina (with local receptive fields and topographically ordered RF centers). Multiple Gaussian spots were presented in each eye as input. Because of cooperation and competition between inputs from the two eyes, groups of neurons developed strong afferent connections to one eye or the other, resulting in patterns of ocular dominance in the network (cf. von der Malsburg 1990; Miller et al. 1989).

The self-organization of the network was studied with varying between-eye correlations. At each input presentation, one spot is randomly placed at (x_i,y_i) in the left retina, and a second spot within a radius of $c \times R$ of (x_i, y_i) in the right retina. The parameter $c \in [0, 1]$ specifies the spatial correlations between spots in the two retinas, and can be adjusted to simulate different degrees of correlations between images in the two eyes. Multi-spot images can be generated by repeating the above step: the simulations below used two-spot images in each eye. Two different simulations are illustrated. In the first one, there were no between-eye correlations in the input ($c = 1$), thus simulating strabismus, where prominent ocular dominance patterns are known to develop. In the second case, there were positive between-eye correlations ($c < 1$), simulating the case of normal binocular vision.

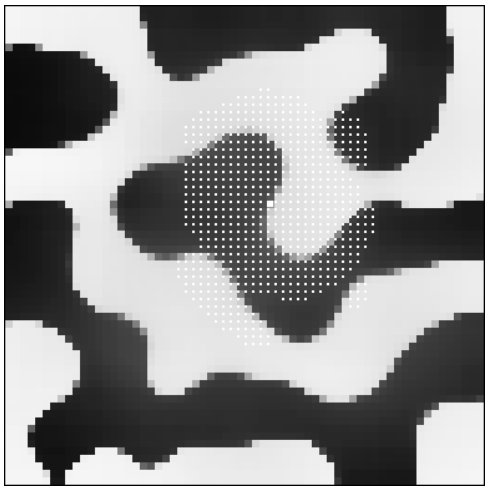
Figure 5 illustrates the response properties and lateral connectivity in the strabismic case. The connection patterns closely follow ocular dominance organization. As neurons become better tuned to one eye or the other, activity correlations between regions tuned to the same eye become stronger, and correlations between opposite eye areas weaker. As a result, monocular neurons develop strong lateral connections to regions with the same eye preference, and weak connections to regions of opposite eye preference. Most neurons are monocular, but a few binocular neurons occur at the boundaries of ocular dominance regions. Because they are equally tuned to the two eyes, the binocular neurons have activity correlations with both ocular dominance regions, and their lateral connection weights are distributed more or less equally between them.

Figure 6 shows the ocular dominance and lateral connection patterns for the normal case. The ocular dominance stripes are narrower and there are more ocular dominance columns in the network. Most neurons are neither purely monocular nor purely binocular and few cells have extreme values of ocular dominance. Accordingly, the lateral connectivity in the network is only partially determined by ocular dominance. How-

[2] Even with large scatter, it is still possible to establish topographic order with additional mechanisms such as chemical markers (Willshaw and von der Malsburg 1979), or special initialization schemes (Willshaw and von der Malsburg 1976). Also, computationally there is no restriction on how large the receptive fields can be. They can cover the whole retina, and global order will develop because such receptive fields overlap a lot (i.e. completely). Of course, anatomically such full connectivity would not be realistic.

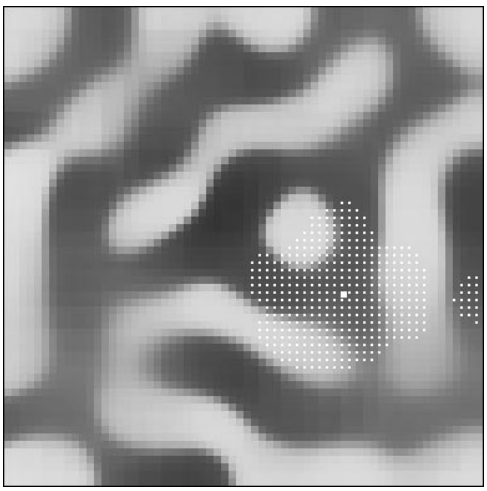


(*a*) Connections of a Monocular Neuron

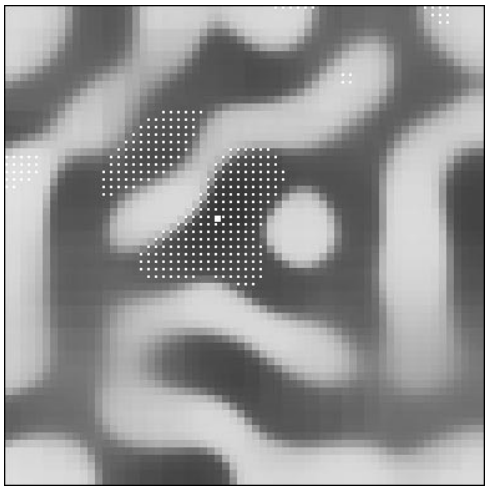


(*b*) Connections of a Binocular Neuron

Figure 5: **Ocular dominance and patterned long-range lateral connections in the strabismic case.** Each neuron is colored with a grey-scale value (*black* $\rightarrow$ *white*) that represents continuously changing eye preference from exclusive left through binocular to exclusive right. Most neurons are monocular, so white and black predominate. Small white dots indicate the strongest lateral input connections to the neuron marked with a big white dot. Only the long range inhibitory connections are shown. The excitatory connections link each neuron only to itself and its eight nearest neighbors. (*a*) The lateral connections of a left monocular neuron predominantly link areas of the same ocular dominance. (*b*) The lateral connections of a binocular neuron come from both eye regions. In this simulation, the parameters were: network size $N = 64$; retinal size $R = 24$; afferent field size $s = 9$; $\delta = 0.1$; $\beta = 0.65$; spot width $a = 5.0$; excitation radius $d = 1$; inhibition radius=31; scaling factors $\gamma_e = 0.5$, $\gamma_i = 0.9$; learning rates $\alpha_{\mathrm{A}} = \alpha_{\mathrm{E}} = \alpha_{\mathrm{I}} = 0.002$; number of training iterations=$35,000$. The anatomical RF centers were slightly scattered around their topographically ordered positions (radius of scatter=0.5, as in figure 2), and all connections were initialized to random weights.



(*a*) Connections of a Neuron with Intermediate Ocularity



(*b*) Connections of a Monocular Neuron

Figure 6: **Ocular dominance and patterned long-range lateral connections in the normal case.** The simulation was otherwise identical to that of figure 5, except that between-eye correlations were greater than zero ($c = 0.4$). The stripes are narrower, and most neurons have intermediate values of ocular dominance (colored gray). Their lateral connections are only partially influenced by ocular dominance, as in (*a*). However, as shown in (*b*), the lateral connections of the most strongly monocular neurons reflect the ocular dominance organization like in the strabismic case. If the between-eye correlations are increased to one (perfectly matched spots in the two retinas), ocular dominance properties will disappear and simple topographic maps and Mexican hat lateral interaction will result, as in figures 2– 4.

ever, the lateral connections of the few strongly monocular neurons follow the ocular dominance patterns like in the strabismic case. In both cases, the spacing between the lateral connection clusters matches the stripe-width.

The patterns of lateral connections and ocular dominance shown above closely match observations in the primary visual cortex. Löwel and Singer (1992) observed that when between-eye correlations were abolished in kittens by surgically induced strabismus, long-range lateral connections primarily linked areas of the same ocular dominance. However, binocular neurons, located between ocular dominance columns, retained connections to both eye regions. The ocular dominance stripes in the strabismics were broad and sharply defined (Löwel 1994). In contrast, ocular dominance stripes in normal animals were narrow and less sharply defined, and lateral connection patterns overall were not significantly influenced by ocular dominance. The receptive field model reproduces these experimental results, and also predicts that the lateral connections of strongly monocular neurons would follow ocular dominance even in the normal case. The model therefore confirms that patterned lateral connections develop based on correlated neuronal activity and demonstrates that they can self-organize cooperatively with ocular dominance columns.

5 Computational Predictions of the Model

RF-LISSOM is a computational model of known physiological phenomena, and it can be used to identify computational constraints that must be met for these phenomena to occur. Below, adaptation and extent of lateral excitation and inhibition, and correlation vs. lateral excitation, are discussed as possible such constraints. Other computational details are included in the appendix.

5.1 Adapting lateral excitation and inhibition

For proper topographic order to develop in our model, the initial lateral excitatory radius must be large enough so that the network produces a single localized spot of activity for each input spot. Otherwise spurious correlations are introduced into the self-organizing process, and the map is likely to develop "twists". A good general guideline is that the initial radius of excitation should be comparable to the range of activity correlations in the network. However, the excitatory radius need not be fixed. As in the Self-Organizing Map algorithm (Kohonen 1982), the radius may start out large, and gradually decrease until it covers the nearest neighbors only. Decreasing the radius in this manner makes the receptive fields gradually narrower and results in finer topographic order. This is a very robust effect and has been observed in numerous experiments.

Even at a fixed excitatory radius, the lateral weights adapt, which has a similar effect as adapting the excitatory radius. If the input correlations are short-range, the lateral excitatory profile will become more peaked, and the inhibition will become stronger in the vicinity of each neuron (figure 4). Therefore, the activity patterns in the network will become smaller and more focused, as happens when the excitatory radius decreases. As a result, the receptive fields will become narrower. For example, small Gaussian input blobs will produce short range activity correlations in the network, and produce smaller receptive fields.

The range of lateral inhibition is not a crucial parameter, as long as it is greater than approximately twice the excitatory radius. If the range is less, activity patterns often split up into several separate bubbles, and distorted maps form. The weak long-range connections may be repeatedly pruned away during self-organization, as in the previous LISSOM model (Sirosh and Miikkulainen 1994). By weight normalization, such pruning will increase the strength of the surviving connections, thereby increasing the inhibition between the correlated areas. As a result, the activity patterns will become better focused, and more specific maps will develop.

5.2 Long-range inhibition vs. excitation

For the self-organization of receptive fields to occur, our model requires that long-range lateral interactions be inhibitory. This is an important prediction of the model. The inhibitory interactions prevent the neural activity from spreading and saturating. As the lateral connections adapt, it is necessary that the lateral inhibition strengthens by correlated activity. If it was weakened instead, inhibition would concentrate on those units that are the least active, in effect self-organizing itself out of the system. The activity patterns would then spread out and saturate, and highly distorted maps would result.

Anatomical studies in the visual cortex show that long-range horizontal connections have mostly excitatory synapses and link primarily to other excitatory neurons (Kisvarday and Eysel 1992; Hirsch and Gilbert 1991; Gilbert et al. 1990). The crucial question is whether afferent stimulation produces mainly excitatory or inhibitory lateral responses at long range. Partial experimental support for long-range inhibition comes from optical imaging studies, which indicate that substantial lateral inhibition exists even at distances of up to 6mm in the primary visual cortex (Grinvald et al. 1994). This range coincides well with the extent of long-range horizontal connections. Other studies such as that of Nelson and Frost (1978) have also reported long-range inhibitory interactions without clearly identifying the anatomical substrate. Due to insufficient evidence, however, the question of whether the primary effect is inhibitory or excitatory remains controversial. Until this issue is resolved and the underlying cortical circuitry is delineated, it will also be difficult to determine if and how lateral inhibition in the cortex strengthens by correlated activity, as predicted by the model.

5.3 Factors affecting the development of ocular dominance

Four main factors influence how ocular dominance columns form: (1) the network size N, (2) the retina size R, (3) the excitation radius d, and (4) the between-eye correlation parameter c. For given values of the network parameters N, R and d, the degree of between-eye correlation determines whether ocular dominance columns form. In simulations with small values of c (closely matched images in both eyes), only topographic maps develop. As c is gradually increased from zero, a sharp transition from a simple topographic map to an ocular dominance map occurs at a threshold $c_t \in (0, 1)$, and stable ocular dominance patterns appear. For example, in the simulations illustrated in figures 5 and 6, $N = 64$, $R = 24$ and $d = 1$, and the threshold at which ocular dominance columns first appear is approximately $c_t = 0.25$. As the network parameters change, c_t appears to change in a fairly simple manner. The threshold is approximately the same for all networks for which the ratio $\frac{dR}{N}$, or the excitatory radius in retinal units, is the same. As this ratio increases (as when the excitation radius increases, or the network becomes smaller), the relative size of activity bubbles on the network becomes larger, increasing correlation and causing transition to ocular dominance at higher levels of c. This is analogous to the self-organizing maps (Kohonen 1982), where an input dimension (i.e. ocular dominance) becomes represented on the map only when the variance in that dimension (c) grows large enough compared to the neighborhood radius ($\frac{dR}{N}$).

Whether the lateral connections adapt or not is not important for the development of ocular dominance. Well-defined stripes can form even if the lateral weights are fixed, although such stripes are broader than those developed with self-organizing lateral connections. Similarly, ocular dominance columns form under a variety of boundary conditions. With self-organized lateral connections, they form at least with sharp and periodic retinal and network boundaries, and with smaller or same-size anatomical RFs at the boundary, although the patterns of the stripes may be qualitatively different in each case.

6 Future Work

Perceptual grouping rules such as continuity of contours, proximity and smoothness appear as long-range activity correlations in the cortex. Von der Malsburg and Singer (1988) and Singer et al.(1990) have suggested that lateral connections store these correlations and perform perceptual grouping and segmentation by synchronizing cortical activity. The Receptive-Field LISSOM model could be extended to test this hypothesis computationally.

So far, we have demonstrated how the inhibitory lateral connections of LISSOM learn to store long-range activity correlations. On the other hand, models such as that of von der Malsburg and Buhmann (1992) show that neuronal firing can be synchronized quite effectively by inhibitory lateral interactions. These two results could be brought together into a LISSOM architecture using neurons with realistic firing properties, and used to study segmentation and binding phenomena.

In future research, we intend to work on self-organization in networks with firing neurons, study what knowledge will be extracted by lateral connections from realistic input, and how this knowledge could be used in feature grouping and segmentation.

7 Conclusion

The Receptive-Field LISSOM model demonstrates that cortical self-organization can be based on short-range excitation, long-range inhibition, and Hebbian weight adaptation. Starting from random-strength connections, neurons develop localized receptive fields and lateral interaction profiles cooperatively and simultaneously. Self-organization takes place with multiple retinal inputs and with multiple anatomical receptive fields of varying sizes, but requires that the receptive fields be large compared to the degree of scatter among their centers. With inputs coming from two retinas at the same time, ocular dominance stripes develop, and lateral connections primarily connect areas of the same ocular dominance. When input correlations between eyes increase, the ocular dominance stripes become narrower, and the influence of ocular dominance on the lateral connection patterns decreases. The model therefore suggests that lateral connections have a dual role in the cortex: to support self-organization of receptive fields, and to learn long-range feature correlations, which may in turn serve as a basis for perceptual grouping and segmentation.

Appendix

Simulation parameters

The main parameters that influence self-organization are the sigmoid activation function, and the strength of the recurrent lateral excitation and inhibition γ_e and γ_i. The activation function (equation 2) affects the self-organizing process in two ways. First, the lower threshold δ determines the minimum input required for a neuron to participate in weight modification. If the weighted sum is below δ, the output activity $\eta_{ij} = 0$, and the weights will not be modified. Only input vectors close enough to the neuron's weight vector cause it to adapt; the higher the δ, the smaller the area of the input space stimulating the neuron. Second, for a given δ, the upper threshold β determines the gain of the neuron activation. If β is close to δ, the neuron is more nonlinear and small differences in its input can produce large changes in its output. When β is lower, the map amplifies smaller differences in the activity pattern. Therefore, the selectivity of receptive fields can be controlled by adjusting δ and β. Good maps are usually obtained for values of δ from 0.0 to 0.2 and β from 0.55 to 0.8. The actual parameters used in each simulation are shown in the figure captions.

For a given sigmoid, if the lateral excitation is too strong (high value of γ_e), saturated activity spots form and distorted maps result. If γ_e is too small, or if γ_i is too large, the activity spots do not become sufficiently focused, and topographic maps do not develop. When the sigmoid slope is in the range 1.0 to 2.0, the excitation radius is $d = 4$ and the inhibition radius is greater than $2d$, the operating ranges are typically $0.8 \leq \gamma_e \leq 1.1$ and $0.7 \leq \gamma_i \leq 1.4$. If the excitation radius is decreased, individual weights will grow stronger because the total lateral excitation weight of each neuron is normalized to 1.0. As a result, the excitation may become too strong. To offset this problem, the value of γ_e will have to decrease with the excitation radius. At the excitation radius of $d = 1$, the operating range of γ_e was $0.4 \leq \gamma_e \leq 0.65$.

Acknowledgments

This research was supported in part by National Science Foundation under grant #IRI-9309273. Computing time for the simulations was provided by the Pittsburgh Supercomputing Center under grants IRI930005P and TRA940029P.

References

Amari, S.-I. (1980). Topographic organization of nerve fields. *Bulletin of Mathematical Biology*, 42:339–364.

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

Dalva, M. B., and Katz, L. C. (1994). Rearrangements of synaptic connections in visual cortex revealed by laser photostimulation. *Science*, 265:255–258.

Gilbert, C. D. (1992). Horizontal integration and cortical dynamics. *Neuron*, 9:1–13.

Gilbert, C. D., Hirsch, J. A., and Wiesel, T. N. (1990). Lateral interactions in visual cortex. In *Cold Spring Harbor Symposia on Quantitative Biology, Volume LV*, 663–677. Cold Spring Harbor Laboratory Press.

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

Grinvald, A., Lieke, E. E., Frostig, R. D., and Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *Journal of Neuroscience*, 14:2545–2568.

Hirsch, J. A., and Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *The Journal of Neuroscience*, 11:1800–1809.

Hubel, D. H., and Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28:229–289.

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

Kisvarday, Z. F., and Eysel, U. T. (1992). Cellular organization of reciprocal patchy networks in layer iii of cat visual cortex (area 17). *Neuroscience*, 46:275–286.

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

Kohonen, T. (1993). Physiological interpretation of the self-organizing map algorithm. *Neural Networks*, 6:895–905.

Löwel, S. (1994). Ocular dominance column development: Strabismus changes the spacing of adjacent columns in cat visual cortex. *Journal of Neuroscience*, 14(12):7451–7468.

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

Miikkulainen, R. (1991). Self-organizing process based on lateral inhibition and synaptic resource redistribution. In *Proceedings of the International Conference on Artificial Neural Networks* (Espoo, Finland), 415–420. Amsterdam; New York: North-Holland.

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

Nelson, J., and Frost, B. (1978). Orientation-selective inhibition from beyond the classical visual receptive field. *Brain Research*, 139:359–365.

Singer, W., Gray, C., Engel, A., König, P., Artola, A., and Bröcher, S. (1990). Formation of cortical cell assemblies. In *Cold Spring Harbor Symposia on Quantitative Biology, Vol. LV*, 939–952. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sirosh, J., and Miikkulainen, R. (1993). How lateral interaction develops in a self-organizing feature map. In *Proceedings of the IEEE International Conference on Neural Networks (San Francisco, CA)*, 1360–1365. Piscataway, NJ: IEEE.

Sirosh, J., and Miikkulainen, R. (1994). Cooperative self-organization of afferent and lateral connections in cortical maps. *Biological Cybernetics*, 71(1):66–78.

Stryker, M., Allman, J., Blakemore, C., Gruel, J., Kaas, J., Merzenich, M., Rakic, P., Singer, W., Stent, G., van der Loos, H., and Wiesel, T. (1988). Group report. Principles of cortical self-organization. In Rakic, P., and Singer, W., editors, *Neurobiology of Neocortex*, 115–136. New York: Wiley.

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

von der Malsburg, C. (1990). Network self-organization. In Zornetzer, S. F., Davis, J. L., and Lau, C., editors, *An Introduction to Neural and Electronic Networks*, chapter 22, 421–432. New York: Academic Press.

von der Malsburg, C., and Buhmann, J. (1992). Sensory segmentation with coupled neural oscillators. vol. 67, 233–242.

von der Malsburg, C., and Singer, W. (1988). Principles of cortical network organization. In Rakic, P., and Singer, W., editors, *Neurobiology of Neocortex*, 69–99. New York: Wiley.

Willshaw, D. J., and von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London B*, 194:431–445.

Willshaw, D. J., and von der Malsburg, C. (1979). A marker induction mechanism for the establishment of ordered neural mappings: its application to the retinotectal problem. *Philosophical Transactions of the Royal Society of London Biol*, 287:203–243.