Self-Organization with Lateral Connections

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

Technical Report AI92-191
December 1992

Abstract

A self-organizing neural network model for the development of afferent and lateral input connections in cortical feature maps is presented. The weight adaptation process is purely activity-dependent, unsupervised, and local. The afferent input weights self-organize into a topological map of the input space. At the same time, the lateral interaction weights develop a smooth "Mexican hat" shaped distribution. Weak lateral connections die off, leaving a pattern of connections that represents the significant long-term correlations of activity on the feature map. The model demonstrates how self-organization can bootstrap itself based on input information only, without global supervision or predetermined lateral interaction. The model can potentially account for experimental observations such as critical periods for self-organization in cortical maps and development of horizontal connections in the primary visual cortex.

1 Introduction

Many areas of the neocortex form two-dimensional topographic maps of peripheral senses. Such maps are believed to develop in a self-organizing process based on cooperation and competition between neurons (Hubel and Wiesel 1965; Miller et al. 1989; von der Malsburg 1973; von der Malsburg and Singer 1988). The Self-Organizing Feature Map (SOFM) algorithm (Kohonen 1989, 1990) is an abstract computational model of this process. At each step during self-organization, the algorithm finds the neuron whose input synaptic weights are closest to the current input vector and changes the input weights of all neurons in its neighborhood towards the input vector. Over time, the input weights self-organize to form a topological map of the input space. Abstracting the possible underlying mechanisms of cooperation and competition, the SOFM model relies on an external supervisor to find the maximally active unit, and invokes an ad-hoc schedule to determine the neighborhood size. Self-organizing feature maps have been applied, for example, into modeling the development of retinotopy, ocular dominance and orientation preference in the visual cortex, and the development of somatotopy in the somatosensory cortex (Obermayer et al. 1990a, 1990b, 1992; Ritter and Schulten 1986).

The neocortex contains an extremely dense network of long range connections tangential to the cortical lamination. These connections may mediate cooperation and competition between neurons during self-organization of afferent connections (Kohonen 1982, 1989; von der Malsburg and Singer

1988). In the visual cortex, the lateral (or horizontal) connections primarily connect areas with similar functional properties, such as iso-orientation columns (Gilbert et al. 1990; Gilbert and Wiesel 1989). These connections develop after birth and attain their specificity through pruning of unused connections (Callaway and Katz 1990, 1991; Luhmann et al. 1986). Lowel and Singer (1992) showed that the survival of intrinsic horizontal connections in the primary visual cortex depends on correlated (i.e. cooccurring) neuronal activity. It is possible that the remaining lateral connections also self-organize based on correlated activation (von der Malsburg and Singer 1988).

Biologically less abstract implementations of self-organizing feature maps have been proposed where the necessary cooperation and competition takes place through excitatory and inhibitory connections between neurons in the map (Kohonen 1989; Miikkulainen 1991). The lateral connections are non-modifiable and distinct from the afferent connections. The connection weight distribution resembles a "Mexican hat": short-range connections are excitatory and long-range connections inhibitory. The lateral interaction is fixed and uniform over the whole network of neurons. How such connectivity would originally come about is not addressed in these models.

In this paper, a computational model for the self-organization of both the afferent and lateral connections of cortical feature maps is presented. The model is called LCSOFM, for Laterally Connected Self-Organizing Feature Map. All connections start out with random weights, and all connections are modified based on correlated activity through Hebbian learning. In the process, the afferent input connections organize into a topological map of the input space, and the lateral connections of each neuron develop a smooth Mexican hat profile. The LCSOFM model gives a computational account for critical periods of self-organization and the development of lateral connections in the cortex. Because self-organization occurs only in regions of correlated activity, such regions may group separately and self-organize independently. This result suggests that the neocortex is a single self-organizing structure where the boundaries of multiple topographical maps are decided during development by correlated activity.

2 The LCSOFM Model

Each neuron in the LCSOFM model has a set of afferent input connections (from the external input to the map) and a set of lateral input connections (from the other neurons in the map). Each neuron develops an initial response as a weighted sum of the activation in its afferent input connections. Lateral interactions between neurons then focus the initial activation pattern into a localized response on the map. After the pattern has stabilized, all connection weights are modified. The process is repeated for a number of external input examples. As the self-organization progresses, the neurons grow more nonlinear and weak connections die off. Below, these general structures and mechanisms of LCSOFM are described in detail. A specific implementation and simulation results are presented in section 3.

2.1 Response Generation

The LCSOFM network is a laminar organization of interconnected neurons (figure 1). Through the afferent connections, every neuron receives the same vector of external input values. The afferent connections are all excitatory. In addition, each neuron has reciprocal excitatory and inhibitory lateral connections with other neurons. Lateral excitatory connections are short-range, connecting only close neighbors in the map. Lateral inhibitory connections run for long distances, and may even implement full connectivity between neurons in the map. Each connection has a characteristic

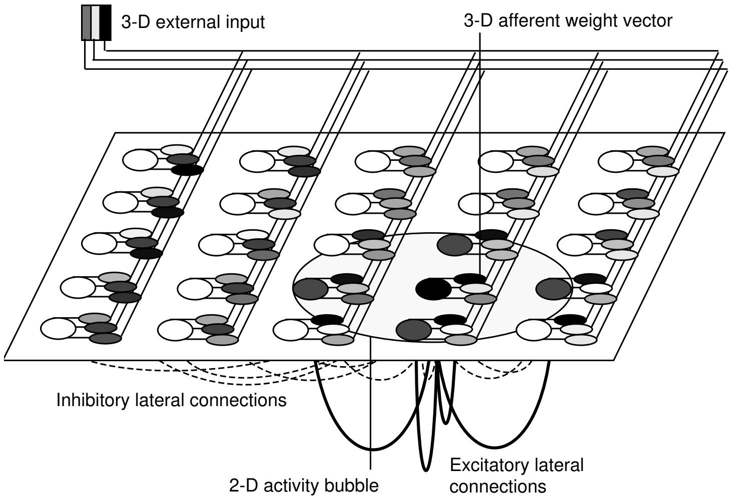


Figure 1: **The laterally connected self-organizing feature map architecture.** Each neuron receives the same 3-dimensional input vector and computes an initial response based on the similarity with its afferent weight vector. The response then evolves through propagation along the lateral connections (only a few connections of the most strongly active unit are shown). After a stable activity bubble has formed, weights of the active units are adapted.

strength (or weight), which may be any value between zero and a prescribed limit that depends on the synaptic resources of the neuron.

Each neuron computes a weighted sum of its inputs and develops a single output activation value by applying a sigmoidal function to the sum. Given a vector of afferent input values $\boldsymbol{\xi}$ and an afferent input weight vector $\boldsymbol{\mu}_{ij}$, the initial response η_{ij} of neuron (i,j) is:

$$\eta_{ij} = \sigma\left(\sum_h \mu_{ij,h}\xi_h\right), \tag{1}$$

where the function σ is a piecewise linear approximation of the familiar sigmoidal activation function (figure 2):

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

The sigmoid introduces a nonlinearity (a soft threshold between δ and β) into the response, and limits the output of the neuron to the range $[0,1]$.

Lateral interaction in the cortex sharpens neuronal response by repeated exchange of activation (Kohonen 1989; Mountcastle 1968). In LCSOFM, the initial response is modified through lateral excitation and inhibition:

$$\eta_{ij}(t) = \sigma\left(\sum_h \mu_{ij,h}\xi_h + \sum_{k,l} E_{kl,ij}\eta_{kl}(t-1) + \sum_{k,l} I_{kl,ij}\eta_{kl}(t-1)\right), \tag{3}$$

where $E_{kl,ij}$ is the excitatory lateral connection weight on the connection from unit (k,l) to unit (i,j), $I_{kl,ij}$ is the inhibitory connection weight, and $\eta_{kl}(t-1)$ is the activity of unit (k,l) during the previous time step. The primary effect of lateral interaction is to sharpen the contrast between areas of high and low activity. If the width of lateral excitation is comparable to the width of the

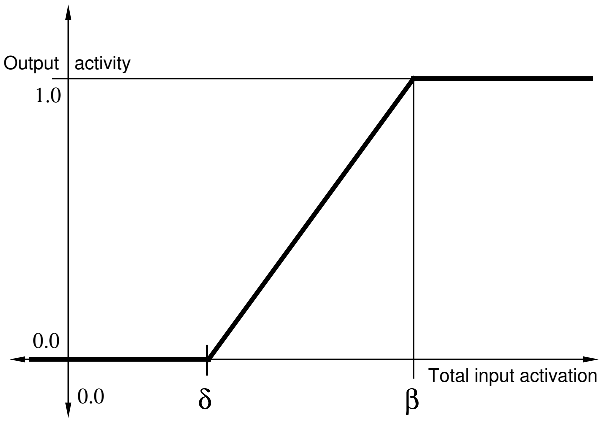


Figure 2: **The neuron's activation function.** The neuron requires an input of at least δ before responding, and saturates at β.

initial response of the network, in successive iterations of equation 3 the response becomes focused around the maximally responding area. In other words, the activity pattern may start out widely spread over a substantial part of the map, but converges iteratively into a stable focused patch of activity, or activity bubble.

Figure 3 illustrates the focusing of the response of an unordered map. The afferent connections are random, and the initial distribution of activity on the map is random also. Repeated exchange of short-range excitation and long-range inhibition enhances activity at the center and suppress activity away from it. If the lateral connectivity is wide enough, the response becomes concentrated to a single cluster close to the center of the map.

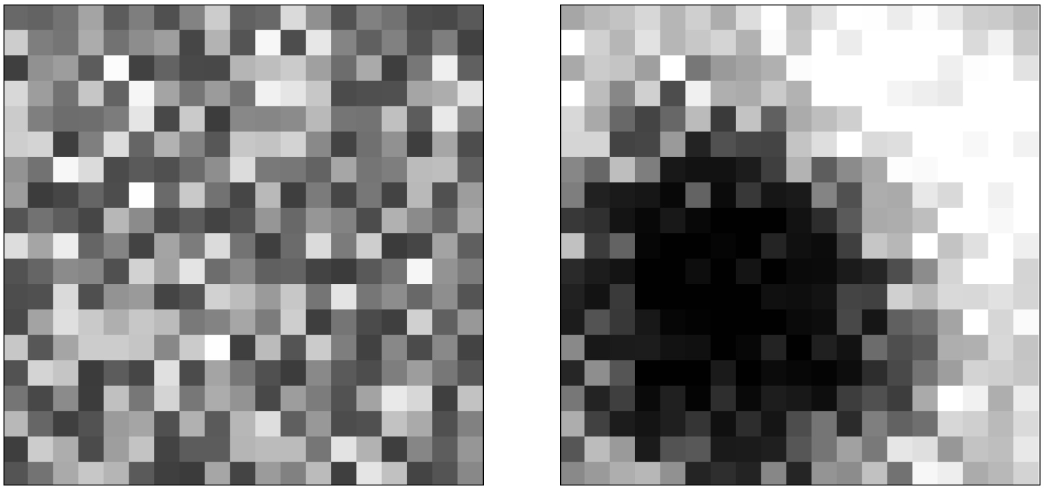
As connection weights in the network become more ordered, the initial activity bubbles become smoother and more regular, and the final stable patterns more focused. Figure 4 shows bubble formation at the center of the map when all the weights are organized. The initial response is a smooth, circular patch of activity. Lateral interactions focus the bubble, brightening the center and diminishing the response at the edges.

2.2 Connection Development

Development in many parts of the brain appears to be guided by Hebbian learning, or similar activity-dependent, correlation-based synaptic adaptation processes (Gustafsson and Wigström 1988; Hebb 1949; Miller and MacKay 1992). There is also evidence that neurons conserve synaptic resources during adaptation (Bourgeois et al. 1989; Hayes and Meyer 1988a, 1988b; Murray et al. 1982; Pallas and Finlay 1991). In LCSOFM, lateral weights develop through a version of the Hebb rule that keeps the total synaptic strength of excitatory and inhibitory lateral connections constant. After the activity on the map has settled into a stable bubble, the lateral connections of unit (i, j) are modified according to:

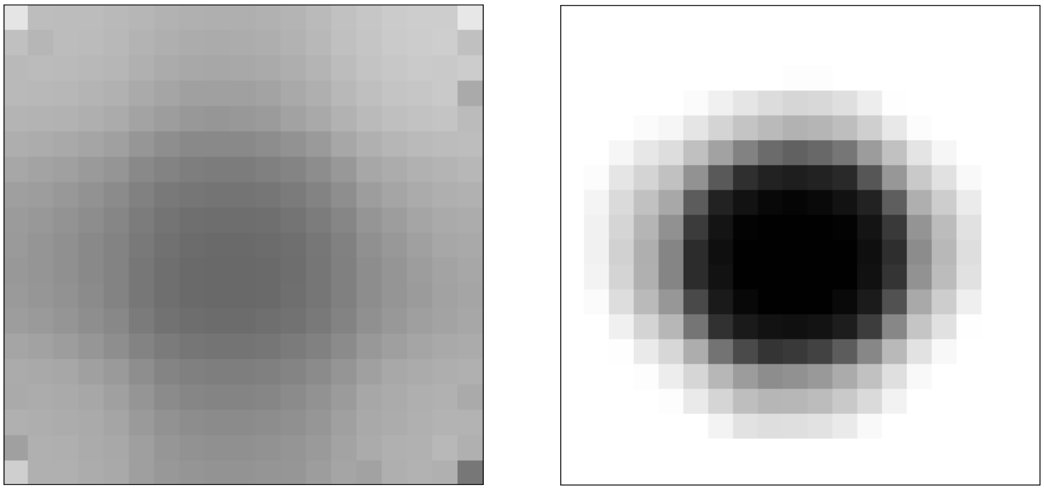
$$\gamma_{ij,kl}(t+1) = \frac{\gamma_{ij,kl}(t) + \alpha_L \eta_{ij} \eta_{kl}}{\sum_{kl} \left[\gamma_{ij,kl}(t) + \alpha_L \eta_{ij} \eta_{kl}\right]}, \tag{4}$$

where $\gamma_{ij,kl}$ is the lateral interaction weight (i.e. $E_{ij,kl}$ or $I_{ij,kl}$), η_{ij} stands for the activity of unit (i, j) in the stable bubble, and α_L is the learning rate (α_E for excitatory weights and α_I for inhibitory). Following the Hebbian principle, the larger the product of pre- and postsynaptic



(a) Initial response (b) Final stable pattern

Figure 3: **Lateral interaction on an unordered map.** Before self-organization, the afferent and lateral connections of the map have random weights, and the initial response to an afferent input is also random (a). On the average the center of the map receives more lateral excitation, and the activation becomes slightly enhanced at the center (b).



(a) Initial response (b) Final stable pattern

Figure 4: **Lateral interaction on an ordered map.** After the map is organized, the initial response to an input at the center of the input space is a smooth, diffuse pattern slightly highlighting the center of the map (a). The final response is a well-focused, high-contrast bubble (b).

activity $\eta_{ij}\eta_{kl}$, the larger the weight change. All connections strengthen by correlated activity; normalization then distributes the changes so that the total weight remains constant.

The afferent input weights are modified according to:

$$\mu_{ij,h}(t+1) = \frac{\mu_{ij,h}(t) + \alpha\eta_{ij}\xi_h}{\left\{\sum_h \left[\mu_{ij,h}(t) + \alpha\eta_{ij}\xi_h\right]^2\right\}^{1/2}}, \tag{5}$$

which is similar to the lateral weight adjustment, but maintains the sum of squares of the afferent weights constant. This is equivalent to redistributing not the total weight itself, but a synaptic resource which is required in quadratic amounts in increasing a synaptic weight. Such normalization is necessary so that the initial response of the neuron to an external input will not be biased by the length of the weight vector.

The above processes of lateral interaction and weight adaptation are sufficient to form ordered afferent and lateral input connections. In order to develop a good feature map of the input space, it is necessary to gradually make these processes more focused and local. As self-organization proceeds in LCSOFM, neuronal responses grow more nonlinear and weak lateral connections die off, which results in more focused activity bubbles and weight changes.

2.3 Modification of Activation Functions

The neuron's activation function (equation 2; figure 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 weights will not be modified (equations 4 and 5). Only input vectors close enough to the neuron's weight vector cause it to adapt; as δ increases, neurons become selective to smaller and smaller areas of the input space.

Second, for a given δ, the upper threshold β determines the gain of the neuron activation. If β is low, the neuron is more nonlinear and small differences in its input can produce large changes in its output. When β decreases, the map amplifies smaller differences in the activity pattern.

Therefore, neurons can be made more selective and sensitive, producing smaller and more refined activity bubbles, by increasing δ and decreasing β. This could be done linearly over time, modeling neuronal growth and maturation uniformly over the map. However, self-organization in LCSOFM does not proceed uniformly. During the early stages, activity bubbles tend to form around the center of the network. Self-organization bootstraps at the center, and gradually spreads towards the periphery of the map. In LCSOFM, δ is increased and β decreased proportional to the activity of the neuron at each input presentation, upto prescribed maximum and minimum limits:

$$\begin{aligned} \delta_{ij}(t+1) &= \min(\delta_{ij}(t) + \alpha_\delta\eta_{ij}, \delta_{\max}), \\ \beta_{ij}(t+1) &= \max(\beta_{ij}(t) + \alpha_\beta\eta_{ij}, \beta_{\min}). \end{aligned} \tag{6}$$

As a result, the neurons grow nonlinear faster at those parts of the map that see more activity. In other words, neurons that work harder mature faster. Such modification automatically takes into account the level of organization around the neuron, which results in more regular final maps.

2.4 Connection Death

During critical periods of cortical development, large numbers of connections die off (Purves and Lichtman 1985). Callaway and Katz (1990) discovered that in the visual cortex, structured horizontal connectivity emerged from an initially unstructured organization after axons projecting to incorrect targets died off. The process appeared to depend on activity-dependent cues in determining which connections should survive.

Along the same lines, connections in the LCSOFM model survive only if they represent significant correlations among neuronal activity. Once the map has organized into a fairly stable state, most of the long-range lateral connections join areas that are no longer ever simultaneously active. Their weights become small, and they can be killed off without disrupting self-organization. Most long-range inhibitory connections are eliminated this way. Since the total synaptic weight is kept constant, inhibition concentrates to the immediate neighborhood of the neuron. The short-range excitatory connections join neurons that are often part of the same bubble. They have fairly large weights and survive.

Before connection death, activity bubbles are large and weights are changed in large neighborhoods. Each adaptation makes neighboring weight vectors more parallel, which means that the map is contracted to the center of the input space. As connections die off, the lateral interaction tends to form smaller, higher-contrast activity bubbles. Weights are changed in smaller neighborhoods, which allows the map to expand and become a better approximation of the input space.

In LCSOFM, a parameter t_O determines the onset of connection death. At t_O, lateral connections with strengths below a threshold w_c are eliminated. From t_O on, more weak connections are eliminated at intervals t_c during the self-organizing process. Eventually, the process reaches an equilibrium where the mapping is stable and all lateral weights stay above w_c.

3 Experiments

In this section, a specific implementation of the LCSOFM model is described, and the process is illustrated in forming a map of a uniform distribution on a 2-dimensional square.

3.1 Input Transformation

In each adaptation step, the afferent weight vector is rotated towards the input vector according to equation 5. The direction of the rotation depends only on the direction of the input vector, not its magnitude (length). The self-organizing process cannot differentiate between inputs in the same direction, and consequently it can only form an $n-1$-dimensional mapping of the n-dimensional input space. However, longer input vectors cause larger changes, and that area of the input space will be magnified in the mapping.

For this reason, the input vectors must be normalized before presenting them to the system. In normalization, the n-dimensional input distribution is projected on the surface of an $n-1$-dimensional hypersphere. Therefore, in order to correctly represent the n-dimensional input space, a redundant $n+1$:th dimension must be added to the input vectors before normalization. The original dimensions are interpreted as angles and the $n+1$:th dimension represents the length of the vector, which is chosen the same for all inputs.

In the simulations below, the original input vectors are uniformly distributed on the 2-dimensional square $x = (x_1, x_2)$, $-0.5 \leq x_1, x_2 \leq 0.5$. Each input vector is transformed into a 3-dimensional cartesian vector $x = (\xi_1, \xi_2, \xi_3)$ whose coordinates in the spherical system are $(x_1, x_2, 1)$:

$$\begin{cases} \xi_1 = 1 \cdot \cos(x_1)\cos(x_2), \\ \xi_2 = 1 \cdot \sin(x_1)\cos(x_2), \\ \xi_3 = 1 \cdot \sin(x_2). \end{cases} \tag{7}$$

When these 3-dimensional input vectors are presented to the LCSOFM model it forms a 2-dimensional mapping of the 3-dimensional input space. Since the dimensions are optimally chosen in the self-organizing process (Kohonen 1989) and the third input dimension is redundant, the mapping directly represents the original input distribution on the 2-dimensional square.

3.2 Simulation Setup

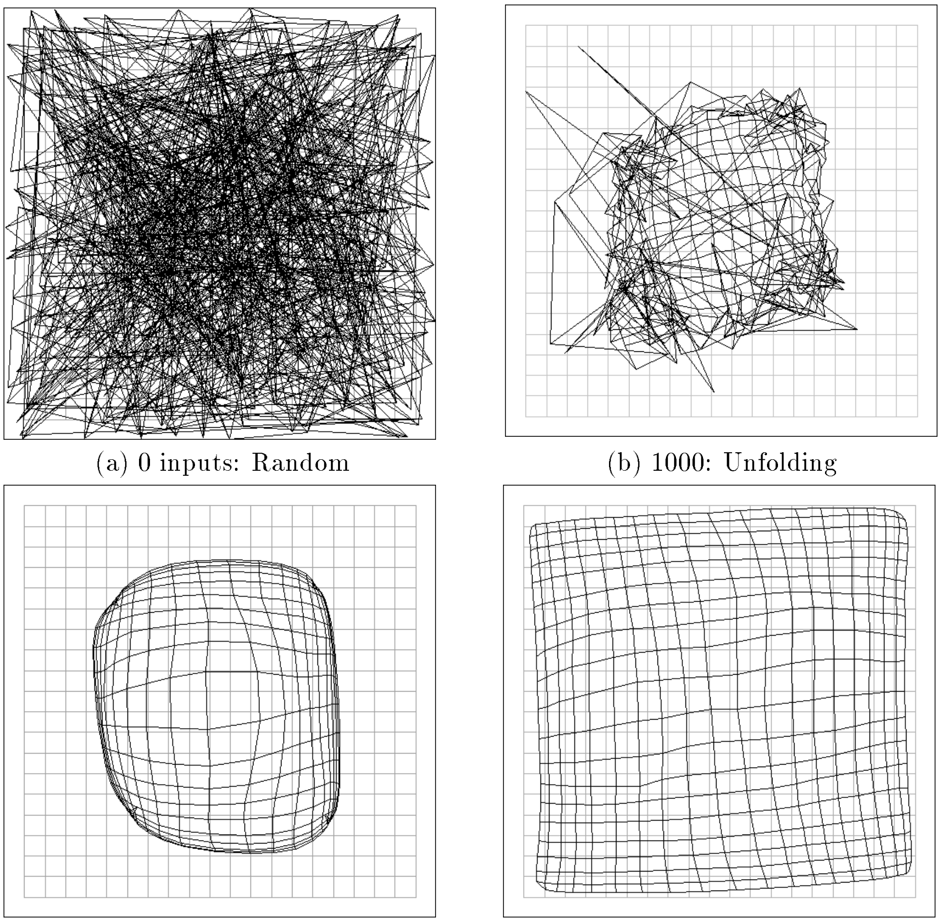
In our implementation of LCSOFM, each neuron initially connects to its neighbors within distance d with excitatory lateral connections and to all neighbors within d' with inhibitory connections (figure 6). Lateral excitation weights are uniformly distributed within $(0,\gamma_e)$ and the inhibition weights within $(0,\gamma_i)$. The parameters d, d', γ_e, and γ_i are chosen so that activity bubbles are regular and comparable to the size of the network diameter. This way the activity can be correlated across the whole map, and an initially random map can be organized.

Experimentally it turns out that a lateral excitation radius $d = 4$ is sufficient for self-organizing a 20×20 LCSOFM network. The lateral inhibition radius d' could be chosen to be any value greater than d; typically, $d' = 3d$ was found adequately large. The initial lateral weights have to be very small so that the neuron responds mostly to the afferent input and self-organization can bootstrap properly. Values $\gamma_e = 0.08$ and $\gamma_i = 0.02$ were found appropriate when the afferent input weights were originally uniformly distributed in the input space. With these weight values, the total input activation of a neuron usually stays within 0.6 and 1.3; these were chosen as the initial sigmoid parameters δ_{ij} and β_{ij}.

3.3 Self-Organization

Self-organization of the LCSOFM network is purely an unsupervised learning process. In our simulation, input items were randomly drawn from the input distribution and presented to the network one at a time. The initial response of each neuron was proportional to the similarity of the input and the afferent weights (equation 1), and this response was refined through lateral interaction (equation 3). After the activity settled into a stable bubble, typically in about 10 iterations, all connection weights were modified according to equations 4 and 5, and the activation function was adapted according to equation 6. The process was repeated for each input vector. After the connection death onset $t_O = 5000$, connections with weights below 0.001 were eliminated every 1000 time steps. The adaptation rates were chosen so that the weights could organize and the activation thresholds δ_{ij} and β_{ij} reach their limits (0.9 and 1.13) before the connections began to die off (i.e. $\alpha = 0.005$, $\alpha_E = \alpha_I = 0.001$, and $\alpha_\delta = \alpha_\beta = 0.00005$).

Figure 5 depicts the evolution of the topological feature mapping. The afferent input weights were transformed back to the spherical coordinates and the original x_1 and x_2 were plotted. The weight vectors are initially uniformly distributed on the input square. As the self-organization progresses, the network unfolds and the weight vectors spread out to form a regular topological map of the input space. Before connection death, the map is stable, but contracted into the center of the input distribution. After connections begin to die off, the map expands and eventually covers the entire input space.



(a) 0 inputs: Random

(b) 1000: Unfolding

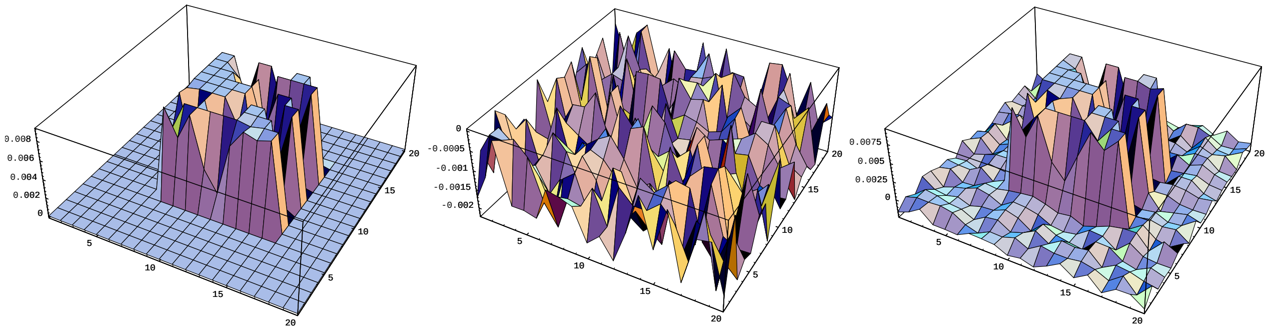
(b) 5,000: Starting connection death

(b) 30,000: Final

Figure 5: **Self-organization of the afferent input weights.** The weight vector of each neuron in the **20 × 20** map is plotted as a point in the original input space (represented by the square). Each weight vector is connected to those of the four immediate neighbors by a line. The resulting dark grid depicts the topological organization of the map. The light grid shows the best possible approximation of the input space by a 20 $times$20 network.

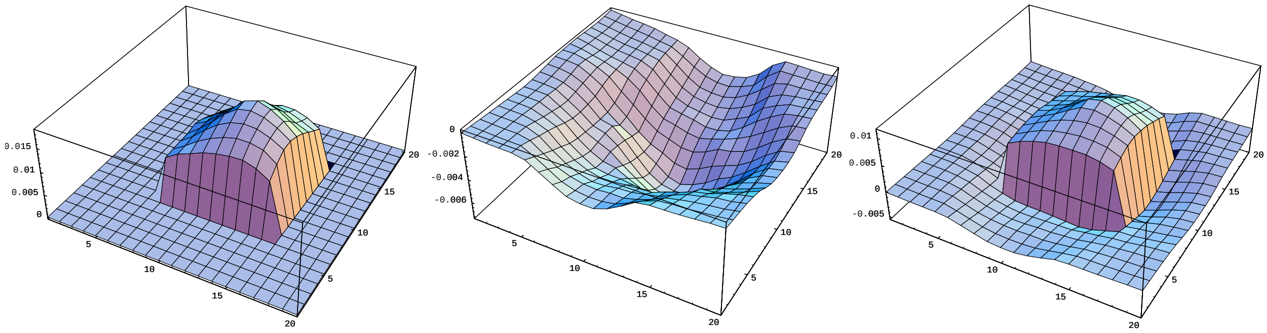
The lateral interconnection weights drive the self-organization of the topological map. They have to converge first before the afferent weights can start to organize. This is exactly what happened in the simulation. By t_O, a very smooth Mexican hat profile had evolved around each neuron (figure 7). After connection death, inhibition became more concentrated, and the combined interaction profiles became lower (figure 8). Such interaction resulted in smaller bubbles, allowing the map to expand.

Interestingly, the lateral interaction profiles are not uniform across the map, but reflect boundary effects (figure 9). Units near the boundary do not have full neighborhoods, and normalization of lateral weights results in profiles that are taller and asymmetric. The neurons at the corner have the tallest interaction profiles. Beyond a distance equal to the excitation radius from the border, the profiles are uniform. The asymmetry is important because it affects the location and shape of the activity bubbles. In simulations with fixed Mexican hat interaction throughout the map, activity bubbles always move away from the boundary even for inputs best stimulating a boundary neuron (Miikkulainen 1991). As a result, the boundaries of the input space are incorrectly represented on



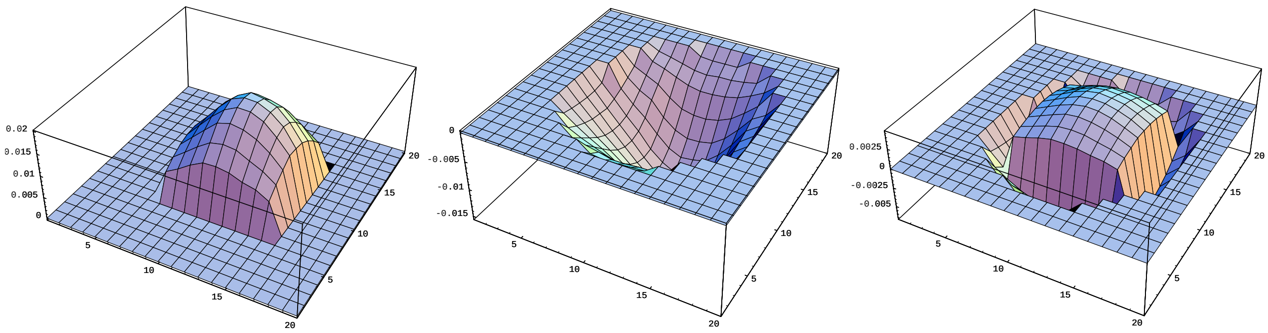
(a) Excitation (b) Inhibition (c) Combined interaction

Figure 6: **Initial lateral interaction.** The lateral excitation and inhibition weights and the combined "Mexican hat" interaction profile are plotted for the neuron at position $(10, 10)$ in the 20×20 map. The excitation weights (a) are initially randomly distributed within the radius $d = 4$, and zero outside. The inhibition weights (b) are randomly distributed within $d' = 12$, and for neuron $(10, 10)$, cover the entire map. The Mexican hat (c) is the sum of the excitatory and inhibitory weights and illustrates the total effect of the lateral connections.



(a) Excitation (b)Inhibition (c) Combined interaction

Figure 7: **Lateral interaction before connection death.** A smooth pattern of excitation and inhibition weights has evolved, resulting in a smooth Mexican hat interaction.



(a) Excitation (b) Inhibition (c) Combined interaction

Figure 8: **Final lateral interaction.** The long-range connections have died off (indicated by zero strength in the figure), resulting in more concentrated and deeper inhibition around the neuron.

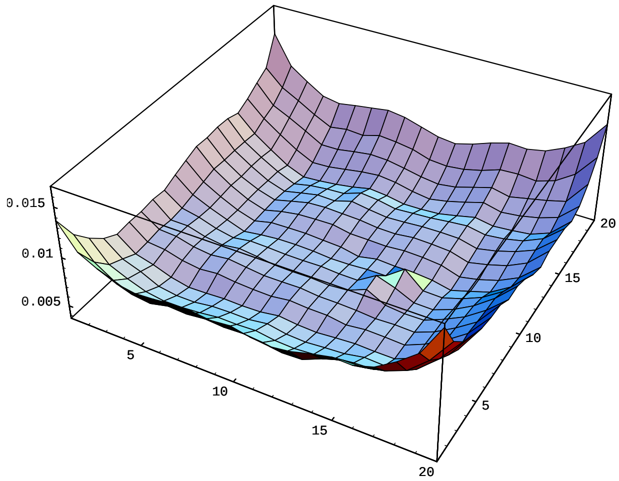


Figure 9: **Distribution of maximum excitation over the map.** The height of the lateral interaction profile is plotted for each neuron in the 20×20 map. Because the lateral weights are normalized, the excitation is higher near the boundary, which allows the activity bubbles to form correctly at the boundary.

the map. In the LCSOFM model, the boundary neurons receive proportionally more excitation, and the bubble forms correctly even at the boundary. In addition, the shape of the bubble resembles the initial response better, and the activity profiles are very smooth. Remarkably, the asymmetries necessary for such behavioor developed automatically as part of the self-organizing process.

4 Discussion and Future Work

The most salient features of the LCSOFM model are its biological basis, its unified mechanism for self-organizing afferent and lateral connections, and its ability to represent non-topographic correlations. As a biological model, LCSOFM provides a computational account for critical periods in cortical self-organization, and is potentially capable of modeling horizontal connection development in the primary visual cortex. Extending LCSOFM into modeling development of multiple feature maps in a single interacting sheet of cortex is a most promising direction for future research.

4.1 How Does LCSOFM Differ from Standard Feature Maps?

LCSOFM is a model of a physical system that self-organizes based on purely local rules. If we were to actually construct "neurons" with properties described above and form a network as explained, the map would self-organize completely in parallel, without global supervision. This is in contrast to the abstract SOFM process, where the maximally responding unit is chosen through global supervision, and adaptation neighborhoods are reduced according to a preset schedule.

The shape of the lateral interaction is automatically extracted from the statistical properties of the external input. At each input presentation, the afferent input vector is analyzed by the current state of the input weights and represented as an activity bubble on the map. The distribution of bubbles over time shapes the lateral weights, which in turn results in tighter and smoother activity bubbles and facilitates the self-organization of the afferent weights. The self-organization thus "bootstraps" by using external input information to establish the necessary cooperative and competive interactions.

Standard feature maps represent the topology of the input space by the network's grid-like layout. Units close by in the map represent input vectors nearby in the input space, and vice versa. In the LCSOFM model, the neighborhood relations are mediated by the lateral connections, and the model is not limited to strictly 2-dimensional topology. Units that are far apart on the map grid may still belong to the same neighborhood if they are strongly connected laterally. Such topologies are automatically learned as part of the self-organizing process. If several areas of the map are simultaneously active, long range connections between such areas can remain strong enough to survive connection death. These connections cause the units to behave as if they were neighbors on the map.

This property of LCSOFM is potentially very significant in representing complex high-dimensional input spaces. While lower-level sensory representation in the brain seems to be organized in 2-dimensional maps (such as retinotopic maps), it is possible that higher representations make use of long-range lateral connections to represent more complex similarity relationships (von der Malsburg and Singer 1988). LCSOFM is a potential computational model for formation of such representations.

4.2 Modeling Cortical Development

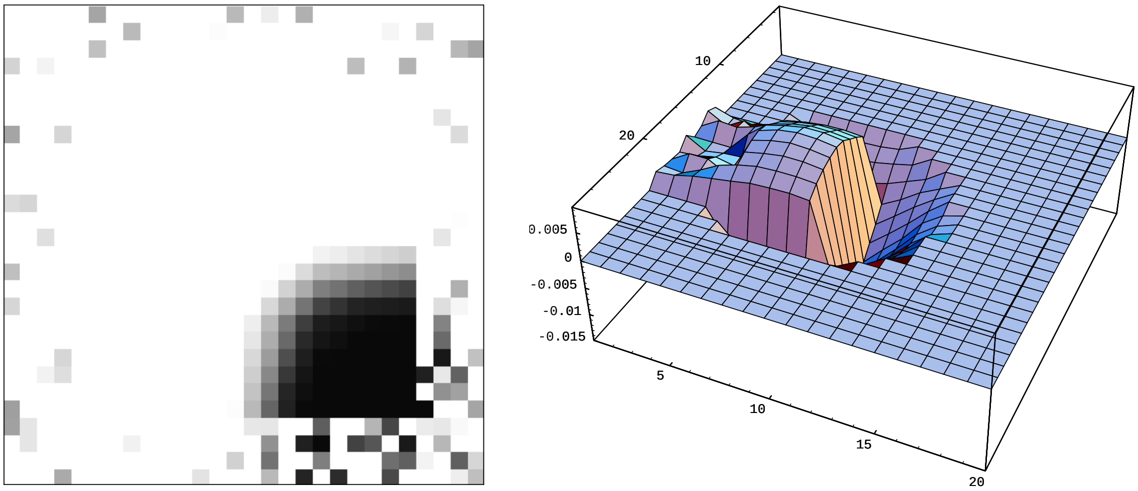
Standard self-organizing feature maps have been used to explain self-organization of afferent input connections to neurons in the primary sensory cortices (Obermayer et al. 1990a, 1990b, 1992). With LCSOFM, it should be possible to model the development of both afferent and lateral input connections in these areas.

It is likely that average visual input to the cortex excites similar feature detectors simultaneously. Given such input, LCSOFM would predict that (1) lateral connections between similar feature detectors strengthen due to correlated activity and (2) connections between dissimilar feature detectors weaken due to limited synaptic resources (through normalization). If weak connections are assumed to die off, the surviving connections would be those that link areas with similar functional properties.

There is evidence for such connectivity in the primary visual cortex (Gilbert and Wiesel 1989; Gilbert et al. 1990). Moreover, it seems that the survival of intrinsic horizontal connections depends on correlated neuronal activity. If visual input to the cortex from both eyes of the cat is decorrelated by artificially introducing strabismus (squint-eye) during development, lateral connections preferentially connect ocular dominance columns of the same eye (Lowel and Singer 1992). In normal cats however, the connections prefer orientation columns of the same orientation specificity.

These results could be explained by the LCSOFM model. In strabismic visual input there would be more correlations among inputs from the same eye than between eyes. Ocular dominance columns representing the same eye would have highly correlated activity and connections between them would strengthen. Other lateral connections would weaken due to normalization and eventually die off. Normally, images in the two eyes are very similar on the average, and have significant correlations. Similar orientation detectors in the cortex would have highly correlated activity irrespective of their ocular dominance. The stronger connections would then run between orientation columns of similar specificity.

In some animals, horizontal connections in the same cortical area run to different lengths in different directions. For example, in areas 17 and 18 of the cat visual cortex, long-range lateral connections align with the iso-orientation slabs. LCSOFM predicts that anisotropic lateral connections are due to anisotropic correlations in activity. Since neurons in an iso-orientation slab respond to



(a) Stable activity pattern (b) Lateral interaction profile

Figure 10: **LCSOFM with uncorrelated surrounding activity.** Only the central 12×12 area of the 20×20 network received input in this experiment. The four outer unit columns and rows were randomly activated. Despite the uncorrelated surroundings, smooth activity bubbles formed in the 12×12 area, even at the corners (a). The lateral connections to the border area mostly decreased to zero and were eliminated. The surviving ones were small and irregular (b).

similar input features, correlations are higher along the length of the slab, and long-range horizontal connections would survive in these directions.

The onset of connection death, t_O, is critical for self-organization in LCSOFM. If connections begin to die too early, some slowly-strengthening connections may be eliminated prematurely, and badly formed maps result. Similarly in the cortex, the onset of connection death seems to be genetically encoded. If the cortex does not receive sufficient input before this time, many relevant connections are eliminated before self-organization is complete. The structure cannot recover unless some of the eliminated connections grow back again (Kandel and Jessell 1991).

4.3 Multiple Self-Organizing Maps

Since LCSOFM networks self-organize according to correlated activity, it might be possible to form separate maps of different input distributions on the same neural sheet. For example, assume that two different input modalities, such as visual and tactile, are fed to different, possibly intersecting regions of a single LCSOFM network. If the activity patterns in the two regions are uncorrelated, the lateral connections linking them will weaken and die. The two regions independently develop separate feature maps. However, if weak correlations existed between the two regions, some of the lateral connections would survive, and implement associations across input modalities.

To demonstrate that self-organization can occur in presence of surrounding uncorrelated activity, a simulation was carried out where the map receiving afferent input was embedded in an area of randomly active neurons (figure 10). The surrounding neurons were connected to the rest of the network with normal lateral connections, and their random initial activity was as strong as the usual initial activation of the neurons in the map. Despite random surroundings, neurons in the map developed coherent activity bubbles. Most of the lateral connections to the randomly-activated neurons died off. A few closest connections survived because the lateral input from the map supported regular bubbles even for neurons near the boundary.

This experiment suggests that the early cortex may well be a uniform, undifferentiated structure with dense lateral connectivity. Differentiation into various topographic maps may be decided dynamically during development based on external input and long-range cooperation and competition. If an area is deprived of input, correlated activity in that region decreases, and other topographic maps may expand into its space. Similarly, if a particular area of the cortex receives intense stimulation, that region of the cortex may expand to occupy a larger area relative to other cortical regions. Such plasticity has indeed been observed in mammalian cortices (Merzenich and Kaas 1982; Merzenich et al. 1984).

5 Conclusion

The LCSOFM model demonstrates how lateral interaction and topological organization of cortical maps can be learned simultaneously from correlations in the input information. The model is biologically motivated, and its predictions agree well with experimental observations on cortical development. LCSOFM is potentially capable of explaining various aspects of lateral and afferent connection development in the cortex, as well as competitive interactions among multiple maps. The model delineates a functional role for connection death and gives a computational account for critical periods in cortical self-organization. These results suggest that a single, general self-organizing process might underlie the development of most aspects of the structure and function of the neocortex.

References

Bourgeois, J. P., Jastreboff, P. J., and Rakic, P. (1989). Synaptogenesis in visual cortex of normal and preterm monkeys: Evidence for intrinsic regulation of synaptic overproduction. *Proceedings of the National Academy of Sciences, USA*, 86:4297–4301.

Callaway, E. M. and Katz, L. C. (1990). Emergence and refinement of clustered horizontal connections in cat striate cortex. *Journal of Neuroscience*, 10:1134–1153.

Callaway, E. M. and Katz, L. C. (1991). Effects of binocular deprivation on the development of clustered horizontal connections in cat striate cortex. *Proceedings of the National Academy of Sciences, USA*, 88:745–749.

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.

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

Gustafsson, B. and Wigström, H. (1988). Physiological mechanisms underlying long-term potentiation. *Trends in Neurosciences*, 11:156–162.

Hayes, W. P. and Meyer, R. L. (1988a). Optic synapse number but not density is constrained during regeneration onto surgically halved tectum in goldfish: HRP-EM evidence that optic fibers compete for fixed numbers of postsynaptic sites on the tectum. *Journal of Computational Neurology*, 274:539–559.

Hayes, W. P. and Meyer, R. L. (1988b). Retinotopically inappropriate synapses of subnormal density formed by misdirected optic fibers in goldfish tectum. *Developmental Brain Research*, 38:304–312.

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory.* New York: Wiley.

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.

Kandel, E. R. and Jessell, T. (1991). Early experience and the fine tuning of synaptic connections. In Kandel, E. R., Schwartz, J. H., and Jessell, T. M., editors, *Principles of Neural Science*, 945–958. Elsevier.

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

Kohonen, T. (1989). *Self-Organization and Associative Memory.* Berlin; Heidelberg; New York: Springer. Third edition.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78:1464–1480.

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

Luhmann, H. J., Millan, M., and Singer, W. (1986). Development of horizontal intrinsic connections in cat striate cortex. *Experimental Brain Research*, 63:443–448.

Merzenich, M. M. and Kaas, J. H. (1982). Reorganization of mammalian somatosensory cortex following peripheral nerve injury. *Trends in Neurosciences*, 5:434–436.

Merzenich, M. M., Nelson, R. J., Stryker, M. P., Cynader, M. S., Schoppmann, A., and Zook, J. M. (1984). Somatosensory cortical map changes following digit amputation in adult monkeys. *Journal of Comparative Neurology*, 224:591–605.

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.

Miller, K. D. and MacKay, D. J. C. (1992). The role of constrains in hebbian learning. CNS Memo 19, Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA.

Mountcastle, V. B. (1968). Neural mechanisms in somesthesia. In Mountcastle, V. B., editor, *Medical Physiology*, vol. 2, 1372–1423. Saint Louis: C. V. Mosby. 12th edition.

Murray, M., Sharma, S., and Edwards, M. A. (1982). Target regulation of synaptic number in the compressed retinotectal projection of goldfish. *Journal of Computational Neurology*, 209:374–385.

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

Obermayer, K., Ritter, H. J., and Schulten, K. J. (1990a). Large-scale simulation of a self-organizing neural network. In *Proceedings of the International Conference on Parallel Processing in Neural Systems and Computers (ICNC)*. New York: Elsevier.

Obermayer, K., Ritter, H. J., and Schulten, K. J. (1990b). A principle for the formation of the spatial structure of cortical feature maps. *Proceedings of the National Academy of Sciences, USA*, 87:8345–8349.

Pallas, S. L. and Finlay, B. L. (1991). Compensation for population-size mismatches in the hamster retinotectal system: Alterations in the organization of retinal projections. *Visual Neuroscience*, 6:271–281.

Purves, D. and Lichtman, J. W. (1985). *Principles of Neural Development*. Sunderland, MA: Sinauer.

Ritter, H. J. and Schulten, K. J. (1986). On the stationary state of Kohonen's self-organizing sensory mapping. *Biological Cybernetics*, 54:99–106.

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

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.