Research Report No.: C-147

Title: A Mathematical Analysis of Development of Oriented Receptive Fields in Linsker's Model

Author: YAMAZAKI Tadashi
Affiliation:
c/o Prof. Osamu Watanabe,
Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology.
email: tyam@is.titech.ac.jp

Date: February 6, 2001

Abstract. We consider Linsker's neural network model, in particular, development of oriented receptive fields of cells on layer G, and investigate which features in this model are crucial for the development. We assume that the correlation function is a Mexican-hat shape function while we introduce the arbor function explicitly and discuss its role in the development. We compute the Fourier transform of the modification rule of synaptic weights to analyze its modification process. Through our mathematical analysis, we obtain the following observations. (1) the Fourier transform of the correlation function extracts a set of Fourier coefficients of a certain frequency, (2) the Fourier transform of the set of those coefficients consists of a finite number of its components, (3) among those components, the arbor function extracts only some of those and others degenerate, (4) thus, the shape of a receptive field can be described as a combination of a certain Fourier cosine components and Fourier sine components, (5) the arbor function determines one of the set of the Fourier cosine components and that of the Fourier sine components dominates the other. We consider two types of the arbor function and analyze when oriented (e.g., bi-lobed) receptive fields are developed. These observations are also justified by computer simulations. Finally we discuss the role of Linsker's parameters k_1 and k_2.

1 Introduction

In a series of papers, Linsker (1986) proposed a multi-layer feed-forward neural network model and studied its self-organization process. He found that cells in lower layers obtain center-surround structured receptive fields while cells in a higher layer, e.g., layer G, obtain bi-lobed (or multi-lobed in general) structured receptive fields[1], by setting parameters of its functions and another additional parameters k_1 and k_2 appropriately through computer simulations .

Inspired by Linsker's works, researchers have investigated Linsker's model from several different view points and have tried to explain how oriented receptive fields are developed. However explanations made so far are not enough to understand it completely. Energy function minimization by Kammen and Yuille (1998) and information theoretic approach by Linsker (1990) do not give clear correspondence between their internal parameters and those in Linsker's model. So it is not clear which properties in Linsker's model are crucial for the development of oriented receptive fields. Principal component analysis by Tang (1989), MacKay and Miller (1990), and fixed point analysis for a given set of weights by Feng, Pan

[1] Linsker called receptive fields with a central excitatory stripe and two inhibitory flanking stripes "bi-lobed". On the other hand, we call them *tri-lobed* throughout this paper, and we call a bi-lobed or a tri-lobed (or multi-lobed in general) receptive field an *oriented* receptive field.

and Roychodwdhury (1997) treat Linsker's model as it is, however, those mainly focus on the development in lower layers. In lower layers, covariance matrices are Gaussian shape while in higher layers the shape is no longer Gaussian. It turns to Mexican-hat shape and therefore the dynamics of development may change. MacKay and Miller (1990) discusses how the leading eigenvector changes in higher layers in section 8 in their paper, however, the observation can not be regarded as a rigorous analysis. In particular, the condition when oriented receptive fields are developed has not been investigated enough. In this paper, we focus on Mexican-hat shape correlation functions and derive some conditions for the development of oriented receptive fields. Then we find that an *arbor function*, a function that determines the synaptic density plays an important role. We show that shapes of receptive fields are affected by the choice of an arbor function and related parameters.

For our analysis, we consider a simplified version of Linsker's model. Our model consists of two neuronal layers that correspond to layer F and layer G in Linsker's model. The lower layer (i.e., layer F) contains a number of cells that project synaptic connections to the upper layer cell, on the other hand, the upper layer (i.e., layer G) consists of a single cell that receives connections from cells in the lower layer. The self-organization process of this model starts by assigning initial synaptic weights between cells in the lower layer and the cell in the upper layer uniformly at random. These weights are updated repeatedly by applying a modification rule, that is, a rule that changes the amount of synaptic weights by computing convolution between the set of synaptic weights and a *correlation function* that describes correlations of activities among cells in the lower layer. In our modification rule, we use a Mexican-hat shape function as the correlation function since in Linsker's model the shape of correlations among cells in layer F was Mexican-hat. Thus, our model is for studying the synaptic weight modification between cells in layer F and a single cell in layer G of Linsker's model. Here we introduce an *arbor function* that determines the synaptic density cells in the lower layer. This function was implicit in Linsker's model, however we consider it explicitly. We consider two types of the arbor function, one is a Gaussian function and the other is a step function.

By using this model, we study which properties and parameter settings are crucial for developing oriented receptive fields by investigating the self-organization process mathematically. We consider the Fourier transform (FT, for short) of our weight modification rule. FT of the correlation function is either an uni-modal function or a bi-modal function. We focus on only the bi-modal case, because the uni-modal case is essentially the same as the case of which the correlation function is a Gaussian function. Since FT of the correlation function is bi-modal, this function extracts only a set of Fourier coefficients of a certain frequency on a circle. (Let σ denote the radius of the circle.) We represent the set of these Fourier coefficients as its Fourier series, that is, we take further FT of the set and decompose it into its Fourier components. Due to the symmetry of signs of Fourier coefficients, only Fourier components that the frequency is a multiple of $2\pi/(4\sigma)$ remain.

Now by examing how these Fourier components affect the synaptic weight modification, we show when and why an oriented receptive field is developed. First we show that only circular symmetric receptive fields are developed if the arbor function is Gaussian. This is because a Gaussian arbor extracts *mainly* the Fourier component of which the frequency is 0. Thus, the set of Fourier coefficients distribute like a flat circle and therefore the shape of the receptive field becomes circular symmetric. Secondly, we consider the case that the arbor is a step function. In this case, the arbor function extracts a set of the Fourier components of which those frequencies are $k2\pi/(4\sigma)$ such that $k = 0, 1, 2, \cdots, 2\alpha - 1$, where α is some constant determined by σ and the range R of the arbor function. Since each component describe the shape of the receptive field, the shape of the receptive field can be computed as a linear combination of the set of these Fourier components. Now for a condition of developing a bi-lobed receptive field, we analyze the case that $\sigma < \pi/R$. In this case, the arbor extracts the Fourier components of frequency 0 and $2\pi/(4\sigma)$. which are respectively a cosine and a sine component. By comparing the magnitude of Fourier sine and cosine parts extracted by the step arbor function, we show that the component of frequency $2\pi/(4\sigma)$ plays a major role in the synaptic growth. Therefore, we show that in our setting, i.e., the arbor is a step function and $\sigma < \pi/R$, a bi-lobed receptive field is almost always developed. (The initial weight only affects the orientation.) We also argue that in the case that several other components are extracted, the shape of receptive fields is affected by the initial weight and is not so robust. These observations are also justified by our computer simulations.

We also discuss the role of parameters k_1 and k_2 in Linsker's model. We point out that these parameters, if they are chosen appropriately, give the same effect as FT of a correlation function becomes bi-modal even if the correlation function itself is a Gaussian function.

This paper is organized as follows. In Section 2 we make a brief overview of Linsker's model and introduce our simplified model and notations of shapes of receptive fields. Section 3 explains our main mathematical results. Here we study the self-organization process of our model analytically. Section 4 is devoted for computer simulations to give justifications for our analysis. In Section 5 we discuss the role of our assumption and that of k_1 and k_2.

2 Preliminaries

In this section, firstly we will make a brief overview of Linsker's model. Since our interest is development of oriented receptive fields, we focus on the afferent synaptic connections between layer F and layer G, where the activity of cells in Layer F follows a Mexican-hat covariance function, and a cell in Layer G obtains an oriented receptive field. Secondly, we consider a correlation based model which is a simplified version of his model and in the following we study instead. Finally, notation of shapes of receptive fields introduced by MacKay and Miller (1990) is reviewed.

Linsker's model

Neural network structure

We consider a multi-layer network model. The model consists of seven neuronal layers called layer $A, B, \ldots, G$, respectively. Each cell of a layer (here we denote as layer $\mathcal{L}$) projects connections approximating a Gaussian distribution centered on a corresponding point in the successive layer (we denote as layer $\mathcal{M}$). It should be noted that Linsker himself introduced lateral connections between cells in layer G. Here however, we omit those connections since the emergence of oriented receptive fields is independent on existing such connections.

Modification rule

The activity level $X^{\mathcal{M}}$ of a cell in layer $\mathcal{M}$ can be computed by the linear sum of its synaptic weights w_i and the activity level of cells in layer L, plus a constant a_1; i.e.

$$X^{\mathcal{M}} = \sum_{i \in \mathcal{L}} w_i X_i^{\mathcal{L}} + a_1.$$

Here it should be noted that $X^{\mathcal{M}}$ might be negative as well as positive. Synaptic weights are modified according to a linear Hebbian rule;

$$\overset{\bullet}{w}_i = c_1 \left(X^{\mathcal{M}} - c_2 \right) \left(X_i^{\mathcal{L}} - c_3 \right) + c_4,$$

where $c_1, \ldots, c_4$ are constants. Synaptic weights are modified following each random presentation of an activity pattern in layer $\mathcal{L}$ to layer $\mathcal{M}$, however assuming that w_i changes on a much longer time scale than the presentations, we can take an average over the ensemble of patterns $X_i^{\mathcal{L}}$ in layer $\mathcal{L}$. Then, the mean rate of change of w_i can be computed as follows.

$$\overset{\bullet}{w}_i = k_1 + \sum_{j \in \mathcal{L}} \left(Q_{ij}^{\mathcal{L}} + k_2 \right) w_j, \tag{1}$$

where k_1, k_2 are constants and $Q_{ij}^{\mathcal{L}} = \left\langle (X_i^{\mathcal{L}} - x^{\mathcal{L}})(X_j^{\mathcal{L}} - x^{\mathcal{L}}) \right\rangle$ is the covariance matrix of points i and j in layer $\mathcal{L}$, and $x^{\mathcal{L}}$ is an average activity of cells in layer $\mathcal{L}$ over the ensemble of presentations. Synaptic weights are constrained by an upper bound and a lower bound, so that $-w_{\text{min}} \leq w_i \leq w_{\text{max}}$ holds.

Computer simulations

Linsker trained this network sequentially, first training the connections from A to B, using random and uncorrelated noise as the inputs from A to B; i.e., Q_{ij}^A is the identity matrix. Once the connections had stabilized, the covariance matrix Q_{ij}^B was computed and used to modify connections between layer B and C. This procedure was repeated for subsequent layers. During this process, the shape of a receptive field showed an ordered structure, in particular, cells in layer G obtained bi-lobed or tri-lobed receptive fields, where the covariance matrix in layer F; i.e., Q_{ij}^F can be expressed as a Mexican-hat function.

Our simplified model

We consider the continuous form of the modification rule, that is, for any cell located at position $\mathbf{x}$ in layer F, we use $w(\mathbf{x})$ instead of w_j to denote its synaptic weight to the cell in layer G. Similarly we denote the covariance between cells at position $\mathbf{x}$ and $\mathbf{y}$ in layer F as $C(\mathbf{x} - \mathbf{y})$, where this is the function of the distance between $\mathbf{x}$ and $\mathbf{y}$, that is, we assume that the covariance between two cells depends on only the distance between those cells. On the other hand, we suppose that an upper bound and a lower bound are large enough and during the process synaptic weights never reach those bounds, therefore we do not take into account the role of those constraints.

In summary, we consider the following equation, which is Eqn. (7) in MacKay and Miller (1990).

$$\overset{\bullet}{w}(\mathbf{x}) = k_1 + \int_{-\infty}^{\infty} \left(C(\mathbf{x} - \mathbf{y}) + k_2 \right) A(\mathbf{y}) w(\mathbf{y}) d\mathbf{y}, \tag{2}$$

$w(\mathbf{x})$ is the synaptic weight of the cell in $\mathbf{x}$ while $\overset{\bullet}{w}(\mathbf{x})$ is the growth rate. $A(\mathbf{x})$ is an *arbor function* that determines the synaptic density of cells in layer F. We consider the following two functions as arbor functions.

$$A(\mathbf{x}) = \begin{cases} 1 & |x| \le R, \\ 0 & \text{otherwise.} \end{cases}, \qquad \text{and} \qquad A(\mathbf{x}) = e^{-\frac{|\mathbf{x}|^2}{2\sigma_A^2}}, \tag{3}$$

where R and $\sqrt{2}\sigma_A$ are positive constants that determine the radius of effective synaptic density, which we call *arbor radius.* We call those functions *Step* and *Gaussian*, respectively.

On the other hand, since Q_{ij}^L can be expressed as a Mexican-hat function, we make use of the following DOG function as our correlation function.

$$C(\mathbf{x}) = K_1 \exp\left(-\frac{|\mathbf{x}|^2}{2\sigma_1^2}\right) - K_2 \exp\left(-\frac{|\mathbf{x}|^2}{2\sigma_2^2}\right), \tag{4}$$

where $\sigma_1 < \sigma_2$. DOG functions of this type are used to express Mexican-hat functions (Miyashita and Tanaka (1992); Miller (1994)).

Initial synaptic weights $w^{(0)}(\mathbf{x})$ are assigned uniformly at random.

Notation of shapes of receptive fields

In order to describe shapes of receptive fields, we make use of the notation introduced by MacKay and Miller (1990). Let $n = 1, 2, 3, \cdots$ denote the total number of radial nodes, and $l =$ s, p, d, $\cdots$ denote the number of angular nodes. For example, '2s' has one radial node and no angular nodes, and '2p' has one angular node and no radial nodes. Figure 1 illustrates these notations.

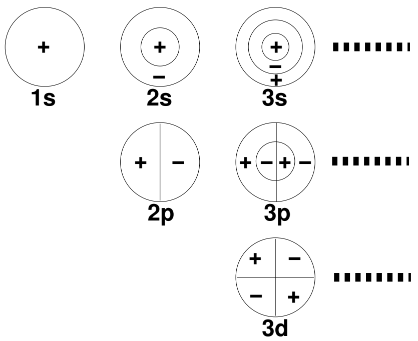

Figure 1: Schematic figures of the notation. $n = 1, 2, 3, \cdots$ denote the total number of nodes and $l =$ s, p, d, $\cdots$ denote the number of angular nodes. 's-modes' are circularly symmetric, 'p-modes' have one angular node, 'd-modes' have two angular nodes, etc.

3 Analytical Results

In our analysis we consider the case of k_1 and k_2 are 0, later we will discuss about the role of these parameters independently.

FT of Eqn.(2) is the following.

$$\overset{\bullet}{\widehat{w}}(\mathbf{n}) = \int_{-\infty}^{\infty} \widehat{A}(\mathbf{n}-\mathbf{m})\widehat{C}(\mathbf{m})\widehat{w}(\mathbf{m})d\mathbf{m},$$

where $\widehat{C}(\mathbf{m})$, FT of the correlation function defined by Eqn.(4), is the following DOG function.

$$\widehat{C}(\mathbf{n}) = 2\pi\sigma_1^2 K_1 \exp\left(-\frac{\sigma_1^2|\mathbf{n}|^2}{2}\right) - 2\pi\sigma_2^2 K_2 \exp\left(-\frac{\sigma_2^2|\mathbf{n}|^2}{2}\right). \tag{5}$$

On the other hand, $\widehat{A}(\mathbf{n})$ is one of the following function.

When A is Gaussian:

$$\widehat{A}(\mathbf{m}) = 2\pi\sigma_A^2 \exp\left(-\frac{\sigma_A^2|\mathbf{m}|^2}{2}\right),$$

When A is Step:

$$\widehat{A}(\mathbf{m}) = 2\pi R \frac{J_1(R|\mathbf{m}|)}{|\mathbf{m}|},$$

where $J_1(x)$ is the Bessel function of the first kind of order 1.

As can be seen from Fig. 2, the shape of $\widehat{C}(\mathbf{n})$ is either uni-modal or bi-modal depending on the choice of parameters. In the following analysis, we consider only the bi-modal case, since the uni-modal case is essentially the same as the case of which the correlation function is Gaussian.

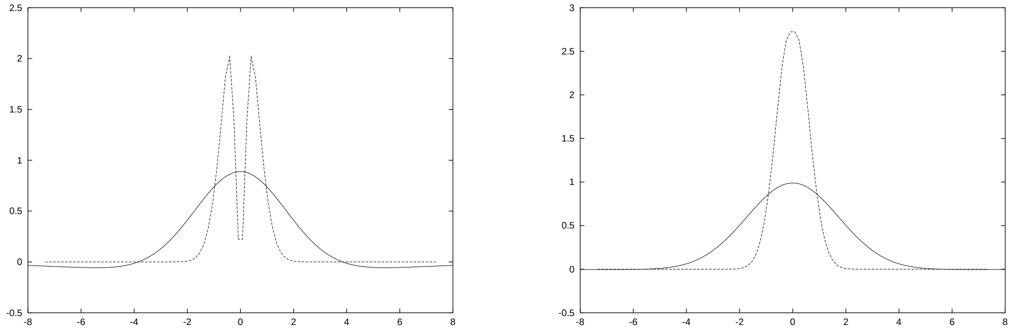

Figure 2: A correlation function (real line) and its FT divided by 2π (dashed line). (Left: under the following parameter setting. $\sigma_1^2 = 8.0$, $\sigma_2^2 = 27.0$,$K_1 = 1.0$, $K_2 = (\sigma_1/\sigma_2)^2$.) (Right: the parameters are the same as the left except $K_2 = 0.001$.)

As shown in Fig. 2 left, $\widehat{C}(\mathbf{n})$ of the bi-modal case can be regarded as follows.

$$\widehat{C}(\mathbf{n}) = \begin{cases} C_*, & \sigma - \varepsilon \leq |\mathbf{n}| \leq \sigma + \varepsilon, \\ 0, & \text{otherwise}, \end{cases}$$

where σ is a constant determined by $K_1, K_2, \sigma_1, \sigma_2$, and $\varepsilon > 0$ is a sufficiently small constant. From this approximation, it follows

$$\dot{\widehat{w}}(\mathbf{n}) = \int_{|\mathbf{m}|\approx\sigma} \widehat{A}(\mathbf{n}-\mathbf{m})\widehat{w}(\mathbf{m})d\mathbf{m}, \tag{6}$$

where $\widehat{A}(\mathbf{m})$ is FT of $A(\mathbf{x})$. Here by $|\mathbf{m}| \approx \sigma$, we mean the integral domain $\sigma - \varepsilon \leq |\mathbf{m}| \leq \sigma + \varepsilon$. This notation will be used for the sake of simplicity. Since the integral domain is limited to the circle such that $|\mathbf{m}| \approx \sigma$, $\widehat{w}(\mathbf{n})$ is effective only when $|\mathbf{n}| \approx \sigma$. Thus we assume here that

$$\widehat{w}(\mathbf{n}) = \begin{cases} \widehat{w}(\mathbf{n}) & |\mathbf{n}| = \sigma, \\ 0 & \text{otherwise.} \end{cases}$$

On the circle such that $|\mathbf{n}| = \sigma$, $\widehat{w}(\mathbf{n})$ can be considered as one-dimensional periodic function. Thus, we introduce polar coordinates,

$$\mathbf{n} \stackrel{\text{def}}{=} \sigma e^{i\theta_n}, \quad \text{and} \quad \mathbf{m} \stackrel{\text{def}}{=} \sigma e^{i\theta_m}.$$

Then we have

$$\begin{aligned} |\mathbf{n} - \mathbf{m}| &= \sqrt{|\mathbf{n}|^2 - 2|\mathbf{n}||\mathbf{m}|\cos(\theta_n - \theta_m) + |\mathbf{m}|^2} \\ &= \sigma\sqrt{2\left(1 - \cos(\theta_n - \theta_m)\right)}. \end{aligned}$$

Therefore its periodicity is 4σ. We denote this one-dimensional periodic function by $\widehat{w}(l)$. Then we decompose this function into its Fourier series as follows.

$$\widehat{w}(l) = \sum_{k=-\infty}^{\infty} \widehat{w}_k(l),$$

where

$$\widehat{w}_k(l) \stackrel{\text{def}}{=} \widehat{\widehat{w}}(k)e^{ikl}.$$

We call each $\widehat{w}_k(l)$ *a Fourier component* of $\widehat{w}(l)$. Due to the symmetry of the sign of each Fourier component such that

$$\Re(\widehat{w}_k(l)) = \Re(\widehat{w}_k(-l)) \quad \text{and} \quad \Im(\widehat{w}_k(l)) = -\Im(\widehat{w}_k(-l)),$$

$\widehat{w}_k(l)$ such that k is a multiple of $2\pi/(4\sigma)$ remain, and the other degenerate. It follows

$$\widehat{w}(l) = \sum_{k=-\infty}^{\infty} \widehat{w}_{k2\pi/(4\sigma)}(l).$$

Here we focus on the shape of $\widehat{w}_{k2\pi/(4\sigma)}(l)$ in the original two-dimensional space, which we denote those by $\widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n})$. Each $\widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n})$ is an oscillating circle. For example, $\widehat{w}_{0\cdot 2\pi/(4\sigma)}(\mathbf{n})$ is a flat circle, $\widehat{w}_{1\cdot 2\pi/(4\sigma)}(\mathbf{n})$ is a slanted circle. Figure. 3 shows that signs of $\widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n})$ for $k = 0, 1, \cdots, 5$. Thus, we can decompose $\widehat{w}(\mathbf{n})$ into the Fourier series as follows.

$$\widehat{w}(\mathbf{n}) = \sum_{k=-\infty}^{\infty} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}).$$

That is, $\widehat{w}(\mathbf{n})$ can be decomposed into a set of oscillating circles $\widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n})$. We also call each $\widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n})$ *a Fourier component* of $\widehat{w}(\mathbf{n})$.

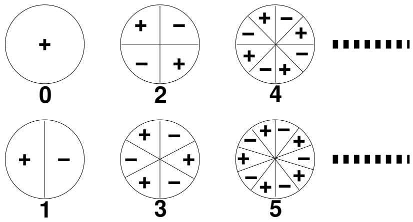

Figure 3: Schematic figures of signs of $\widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n})$. For each figure, an integer below the figure represents k, and '+' and '−' denote the signs of $\widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n})$ on the circle such that $|\mathbf{n}| = \sigma$. It should be noted that k corresponds to the number of angular nodes which is denoted by s,p,d, $\cdots$ in MacKay and Miller (1990).

Among these components of $\widehat{w}(\mathbf{n})$, only some of them are extracted by $A(\mathbf{x})$, because FT of Eqn. (6) is

$$\dot{\widehat{\widehat{w}}}(\mathbf{u}) = \widehat{\widehat{A}}(\mathbf{u})\widehat{\widehat{w}}(\mathbf{u}),$$

and its solution is as follows.

$$\widehat{\widehat{w}}(\mathbf{u}) = \widehat{\widehat{w}}^{(0)}(\mathbf{u}) \exp\left(\widehat{\widehat{A}}(\mathbf{u})t\right).$$

In the following, we consider which Fourier components of $\widehat{w}(\mathbf{n})$ are extracted by two different arbor functions and what are the receptive fields made by these extracted Fourier components. Firstly we discuss the case that $A(\mathbf{x})$ is Gaussian, then the case that $A(\mathbf{x})$ is Step is studied.

In the case of which $A(\mathbf{x})$ is Gaussian, $\widehat{\widehat{A}}(\mathbf{u})$ is also a Gaussian. Since $\widehat{\widehat{A}}(\mathbf{u})$ shows its maximum value on $\mathbf{u} = 0$, $\exp(\widehat{\widehat{A}}(\mathbf{u})t)$ extracts mainly $\widehat{w}_0(\mathbf{n})$. Thus, $\widehat{w}(\mathbf{n})$ can be written as

$$\widehat{w}(\mathbf{n}) \approx \widehat{w}_0(\mathbf{n}) \stackrel{\text{def}}{=} c,$$

where c is some constant. Therefore it follows

$$w(\mathbf{x}) = \int\limits_{|\mathbf{n}|\approx\sigma} \widehat{w}(\mathbf{n}) \cos(\mathbf{n}\mathbf{x}) d\mathbf{n} = c \int\limits_{|\mathbf{n}|\approx\sigma} \cos(\mathbf{n}\mathbf{x}) d\mathbf{n}.$$

The shape is a concentric circle like 's-mode' and the number of radial nodes is in proportional to the value of σ.

In contrast to Gaussian, if A is Step, since $\widehat{\widehat{A}}(\mathbf{u})$ is also a step function of the range $[-R, R]$, $\widehat{\widehat{A}}(\mathbf{u})$ extracts a set of frequencies from the range $[-R, R]$. For some α, if there is a relation between σ and R as follows,

$$\sigma < \frac{\alpha\pi}{R},$$

we have

$$\widehat{w}(\mathbf{n}) = \sum_{k=-2\alpha-1}^{2\alpha-1} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}),$$

because

$$\frac{m2\pi}{4\sigma} < R \Longleftrightarrow m < 2\alpha.$$

For example, in the case of $\alpha = 1$, that is, $\sigma < \pi/R$,

$$\widehat{w}(\mathbf{n}) = \widehat{w}_0(\mathbf{n}) + \widehat{w}_{2\pi/(4\sigma)}(\mathbf{n}).$$

(It should be noted that the condition $\sigma < \pi/R$ is crucial. If we relax the condition and increase σ a bit, then the Fourier component $\widehat{w}_{2\cdot 2\pi/(4\sigma)}(\mathbf{n})$ is also extracted and therefore the shape of receptive fields changes.)

We decompose $w(\mathbf{x})$ into the Fourier cosine part and the Fourier sine part as follows.

$$w(\mathbf{x}) = w_{\cos}(\mathbf{x}) + w_{\sin}(\mathbf{x}),$$

where

$$\begin{aligned} w_{\cos}(\mathbf{x}) &= \sum_{k \text{ is even.}} \int_{|\mathbf{n}|\approx\sigma} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}) \cos(\mathbf{n}\mathbf{x}), \\ w_{\sin}(\mathbf{x}) &= \sum_{k \text{ is odd.}} \int_{|\mathbf{n}|\approx\sigma} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}) \sin(\mathbf{n}\mathbf{x}). \end{aligned} \tag{7}$$

Here we notice that the Fourier components of $\widehat{w}(\mathbf{n})$ can be separated into the Fourier cosine components and the Fourier sine components. For any $k = 0, 1, 2, \cdots$, $\widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n})$ is a Fourier cosine component of $\widehat{w}(\mathbf{n})$ if k is even while it is a Fourier sine component if k is odd (See Fig. 3). Thus, we have the above decomposition.

In the following we consider the case of $\alpha = 1$, that is $\sigma < \pi/R$, since this is the case of developing a stable oriented receptive field, that is, a bi-lobed receptive field. Later we discuss a more general case for any α.

Eqn. (7) can be described as follows.

$$\begin{aligned} w_{\cos}(\mathbf{x}) &= \int_{|\mathbf{n}|<\pi/R} \widehat{w}_0(\mathbf{n}) \cos(\mathbf{n}\mathbf{x}), \\ w_{\sin}(\mathbf{x}) &= \int_{|\mathbf{n}|<\pi/R} \widehat{w}_{2\pi/(4\sigma)}(\mathbf{n}) \sin(\mathbf{n}\mathbf{x}). \end{aligned}$$

In the same way of which $A(\mathbf{x})$ is Gaussian, $w_{\cos}(\mathbf{x})$ becomes a concentric circle like '1s' or '2s'. On the other hand, since $\widehat{w}_{2\pi/(4\sigma)}(\mathbf{n})$ takes positive values on a semi circle and negative values on the other semi circle (see Fig. 3), $\widehat{w}_{2\pi/(4\sigma)}(\mathbf{n})$ is a slanted circle. Therefore there must be the frequency $\mathbf{n}_*$ such that

$$\widehat{w}_{2\pi/(4\sigma)}(\mathbf{n}_*) = \max_{\mathbf{n}}\{\widehat{w}_{2\pi/(4\sigma)}(\mathbf{n})\}.$$

Thus, the sine components in the both sides of $\mathbf{n}_*$ are canceled out and $\sin(\mathbf{n}_*\mathbf{x})$ remains. Hence $w_{\sin}(\mathbf{x})$ becomes bi-lobed like '2p' (Notice that $|\mathbf{n}_*| = \pi/R$).

Finally we determine which one of $w_{\cos}(\mathbf{x})$ and $w_{\sin}(\mathbf{x})$ dominates the other. We decompose Eqn. (6) into the Fourier cosine part and the Fourier sine part as follows.

$$\dot{\widehat{w}}(\mathbf{n}) = \int\limits_{|\mathbf{m}|\approx\sigma} \widehat{A}_{\cos}(\mathbf{n},\mathbf{m})\widehat{w}_{\cos}(\mathbf{m})d\mathbf{m} + \int\limits_{|\mathbf{m}|\approx\sigma} \widehat{A}_{\sin}(\mathbf{n},\mathbf{m})\widehat{w}_{\sin}(\mathbf{m})d\mathbf{m}, \tag{8}$$

where

$$\widehat{A}_{\cos}(\mathbf{n},\mathbf{m}) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} A(\mathbf{x})\cos(\mathbf{n}\mathbf{x})\cos(\mathbf{m}\mathbf{x})d\mathbf{x},$$
$$\widehat{A}_{\sin}(\mathbf{n},\mathbf{m}) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} A(\mathbf{x})\sin(\mathbf{n}\mathbf{x})\sin(\mathbf{m}\mathbf{x})d\mathbf{x},$$

$$\widehat{w}_{\cos}(\mathbf{n}) \stackrel{\text{def}}{=} \sum_{k \text{ is even.}} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}),$$
$$\widehat{w}_{\sin}(\mathbf{n}) \stackrel{\text{def}}{=} \sum_{k \text{ is odd.}} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}).$$

(Notice that due to the symmetry of the Fourier coefficients by the definition of FT, it is sufficient to consider only the first and second quadrants as the integral domain, although we consider all quadrants in the above integrals.)

We will observe that the magnitude of $\widehat{A}_{\cos}(\mathbf{n},\mathbf{m})$ and that of $\widehat{A}_{\sin}(\mathbf{n},\mathbf{m})$ in Eqn. (8) are different, which we denote by $\widehat{A}_{\cos}$ and $\widehat{A}_{\sin}$ respectively. In other words, one of $w_{\cos}(\mathbf{x})$ and $w_{\sin}(\mathbf{x})$ appears as the shape of the receptive field. Since $\widehat{A}_{\cos}(\mathbf{n},\mathbf{m})$ and $\widehat{A}_{\sin}(\mathbf{n},\mathbf{m})$ are circular symmetric with respect to $\mathbf{n}$ such that $|\mathbf{n}| \approx \sigma$, we fix $\mathbf{n}$ as $(\sigma, 0)$. Then we compute those magnitudes $\widehat{A}_{\cos}$ and $\widehat{A}_{\sin}$ as follows.

$$\widehat{A}_{\cos} \stackrel{\text{def}}{=} \int\limits_{|\mathbf{m}|\approx\sigma} \widehat{A}_{\cos}(\mathbf{n},\mathbf{m})d\mathbf{m},\Big|_{\mathbf{n}\approx(\sigma,0)} = \int\limits_{|\mathbf{m}|\approx\sigma}\int_{-\infty}^{\infty} A(\mathbf{x})\cos(\mathbf{n}\mathbf{x})\cos(\mathbf{m}\mathbf{x})d\mathbf{x}d\mathbf{m},\Big|_{\mathbf{n}\approx(\sigma,0)},$$
$$\widehat{A}_{\sin} \stackrel{\text{def}}{=} \int\limits_{|\mathbf{m}|\approx\sigma} \widehat{A}_{\sin}(\mathbf{n},\mathbf{m})d\mathbf{m},\Big|_{\mathbf{n}\approx(\sigma,0)} = \int\limits_{|\mathbf{m}|\approx\sigma}\int_{-\infty}^{\infty} A(\mathbf{x})\sin(\mathbf{n}\mathbf{x})\sin(\mathbf{m}\mathbf{x})d\mathbf{x}d\mathbf{m},\Big|_{\mathbf{n}\approx(\sigma,0)},$$

Unfortunately we can not compute $\widehat{A}_{\cos}$ and $\widehat{A}_{\sin}$ analytically. Thus we consider an approximation. Since $\widehat{A}_{\cos}(\mathbf{n},\mathbf{m})$ and $\widehat{A}_{\sin}(\mathbf{n},\mathbf{m})$ take the maximum values when $\mathbf{n} = \mathbf{m}$, we consider only the case of $\mathbf{n} = \mathbf{m}$. Then it follows,

$$\widehat{A}_{\cos}(\mathbf{n},\mathbf{n}) = \int_{-\infty}^{\infty} A(\mathbf{x})\cos^2(\mathbf{n}\mathbf{x})d\mathbf{x}, \quad \text{and} \quad \widehat{A}_{\sin}(\mathbf{n},\mathbf{n}) = \int_{-\infty}^{\infty} A(\mathbf{x})\sin^2(\mathbf{n}\mathbf{x})d\mathbf{x}.$$

Here we can observe that $\widehat{A}_{\cos}(\mathbf{n},\mathbf{n})$ is less than $\widehat{A}_{\sin}(\mathbf{n},\mathbf{n})$ (Fig. 4). Hence, in this case $w_{\sin}(\mathbf{x})$ dominates $w_{\cos}(\mathbf{x})$, and therefore a bi-lobed receptive field like '2p' is developed.

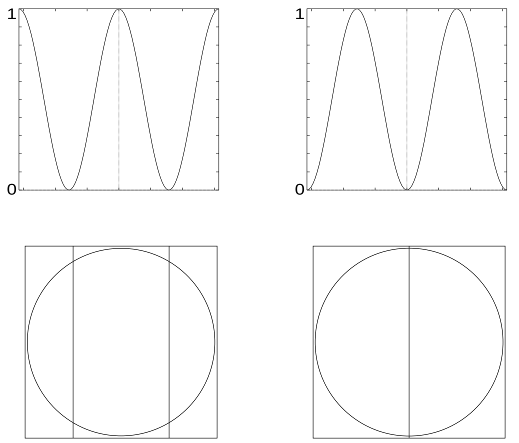

Figure 4: One-dimensional representation of $\cos^2(\mathbf{nx})$ of the fundamental frequency with respect to the horizontal axis (top left) and that of $\sin^2(\mathbf{nx})$ (top right). The peak value of $\cos^2(\mathbf{nx})$ appears at the origin and the both borders of the horizontal axis, while that of $\sin^2(\mathbf{nx})$ appears at the midpoint between the origin and the border. Applying $A(\mathbf{x})$ to $\cos^2(\mathbf{nx})$ eliminates the value at the border, so $\widehat{A}_{\cos}(\mathbf{n},\mathbf{n})$ becomes much smaller than that before applying $A(\mathbf{x})$. On the other hand, $A(\mathbf{x})$ does not affect so much to $\widehat{A}_{\sin}(\mathbf{n},\mathbf{n})$ since at the border the value of $\sin^2(\mathbf{nx})$ is small enough.

Finally let us consider a more general case when α is some integer greater than 1. It follows.

$$\widehat{w}_{\cos}(\mathbf{n}) = \sum_{k=0,2,\cdots,2(\alpha-1)} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}),$$
$$\widehat{w}_{\sin}(\mathbf{n}) = \sum_{k=1,3,\cdots,2\alpha-1} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}).$$

Since one of $\widehat{w}_{\cos}(\mathbf{n})$ and $\widehat{w}_{\sin}(\mathbf{n})$ may dominate, the shape of the receptive field is one of

$$w_{\cos}(\mathbf{x}) = \sum_{k=0,2,\cdots,2(\alpha-1)} \int_{|\mathbf{n}|\approx\sigma} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}) \cos(\mathbf{nx}), \quad \text{and}$$
$$w_{\sin}(\mathbf{x}) = \sum_{k=1,3,\cdots,2\alpha-1} \int_{|\mathbf{n}|\approx\sigma} \widehat{w}_{k2\pi/(4\sigma)}(\mathbf{n}) \sin(\mathbf{nx}).$$

It should be noticed that the relative difference between two sine (resp., cosine) components does not change. Thus, the shape of the receptive field is determined by the initial combination of the sine (resp., cosine) components. For example, if $\alpha = 2$ and if the cosine components become major, then the shape is determined by the $\widehat{w}_0(\mathbf{n})$ and $\widehat{w}_{2\cdot 2\pi/(4\sigma)}(\mathbf{n})$, and their relation is determined by the randomly chosen initial weight. (Note that $\widehat{w}_{2\cdot 2\pi/(4\sigma)}(\mathbf{n})$ develops a '3d' shape field.)

4 Computer Simulations

In this section, we perform computer simulations to give justifications to our analytical results. We consider the case of $\sigma < \pi/R$ to see that either the Fourier cosine part and the Fourier sine part dominates the other determined by the arbor function as well as only Fourier components of certain frequencies are extracted. Firstly we show the setting of our computer simulations and then we see the results.

Neural network structure

We consider two layer neural network. The lower layer is represented by 13 by 13 grids of cells while the upper layer consists of one cell. The cell on the upper layer connects to a circle of diameter 13 grid intervals of cortical cells, in other words, the arbor radius is 6.5.

Modification rule

We make use of Eqn. (2), (3), (4) with the following parameters. For the correlation function, they are $\sigma_1^2 = 3.0$,$\sigma_2^2 = 27.0$, $K_1 = 1.00$, $K_2 = (\sigma_1/\sigma_2)^2 = 0.111$. Hence $\sigma \approx 0.46$. With this setting the shape of the correlation function is represented as Fig. 2 left. On the other hand, for the arbor functions, $R = 6.5$ and $\sqrt{2}\sigma_A = 4.0$ are used. Since $\pi/R \approx 0.48$, $\sigma < \pi/R$ holds for the step arbor function case.

Self-organization process

The initial synaptic weights are assigned from $[-0.4, 0.4]$ uniformly at random. As the pseudo random generator to assign initial synaptic weights, we use MT (Matsumoto and Nishimura (1998)). Then the modification rule is applied in steps. After 80 steps, we stop the self-organization process and compute the shape of the receptive field.

In following, we investigate two computer simulations for each arbor function.

The case of which A is Step

Firstly we compute $A^{\cos}$ and $A^{\sin}$ to predict what kind of receptive fields are obtained with this setting. Since a calculation shows that $A^{\cos}/A^{\sin} \approx 0.70$, we expect that with this setting, a bi-lobed receptive field would be obtained.

The following is the result of a computer simulation. Figure 5 consists of from left to right the distribution of initial synaptic weights, the final shape of the receptive field, its Fourier cosine part and its sine part. respectively. As can be seen, although the initial

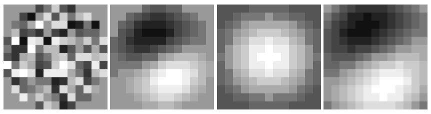

Figure 5: (From left to right) the distribution of initial synaptic weights, the final shape of the receptive field, its Fourier cosine part and its sine part In each figure, weights are normalized.

synaptic weights are assigned uniformly at random, after a number of steps, a bi-lobed receptive field is developed. Moreover, the final shape of the receptive field is as the same as that of its Fourier sine part while its Fourier cosine part becomes a concentric circle.

Figure 6 shows the values of the Fourier cosine and sine coefficients in the first five steps. We can see that, at the beginning of the simulation, both of the Fourier coefficients oscillate. After that, those oscillations disappear. Finally, the circle of the Fourier cosine coefficients is dominated by $\widehat{w}_0(\mathbf{n})$ and arranged to be flat, on the other hand, the circle of the sine coefficients is dominated by $\widehat{w}_{2\pi/(4\sigma)}(\mathbf{n})$ and becomes slanted, and only one coefficient shows the maximum value.

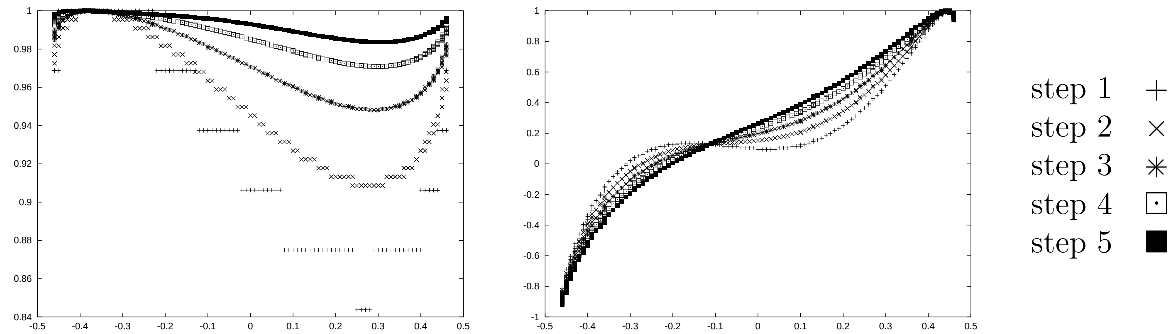

Figure 6: The growth of the Fourier cosine coefficients (left) and the sine coefficients (right) in the first five steps. The horizontal axis shows frequencies of Fourier cosine and sine coefficients while the vertical axis does values of those cosine and sine coefficients. Values of coefficients in each step are normalized and plotted. We plot only the coefficients in the first and second quadrants of the Fourier space. As can be seen, at the first step, the set of those coefficients oscillates, on the other hand, as the self-organization proceeds by, those oscillations disappear.

The case of which A is Gaussian

In this case, $\widehat{w}_0(\mathbf{n})$ dominates and therefore a circular symmetric, in particular, '2s' type in this case, receptive field will be developed. Figure 7 shows from left to right, the final shape of the receptive field, its Fourier cosine and sine part. As can be seen, a circular symmetric receptive field is developed as we expected. Since in our setting, $\pi/(2R) = 0.24$ is small enough, the Fourier component of frequency $\widehat{w}_{2\pi/(4\sigma)}(\mathbf{n})$ is extracted a little. Therefore the Fourier sine part appears.

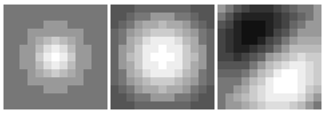

Figure 7: (From left to right) the final shape of the receptive field, its Fourier cosine part and sine part.

5 The role of k_1, k_2

So far we saw that the bi-modal shape of $\widehat{C}(\mathbf{n})$ is crucial. Here we will see that good choice of k_1, k_2 gives similar effects even in the case of $\widehat{C}(\mathbf{n})$ is uni-modal (e.g., Gaussian). MacKay and Miller (1990) have also obtained similar results.

Firstly we discuss the role of k_1. As we saw, our modification rule (Eqn. (2)) can be expressed in terms of FT as follows.

$$\overset{\bullet}{w}(\mathbf{x}) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\widehat{A}(\mathbf{n}-\mathbf{m})\widehat{C}(\mathbf{m})\widehat{w}(\mathbf{m})e^{i\mathbf{n}\mathbf{x}}d\mathbf{m}d\mathbf{n}.$$

The following equation is identical to the above one.

$$\overset{\bullet}{w}(\mathbf{x}) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\widehat{C}(\mathbf{n})\int_{-\infty}^{\infty}\widehat{A}(\mathbf{n}-\mathbf{m})\widehat{w}(\mathbf{m})e^{i\mathbf{n}\mathbf{x}}d\mathbf{m}d\mathbf{n}.$$

We separate the right hand side into the component that $\mathbf{n} = \mathbf{0}$ and the others. Then the rule turns into the following one.

$$\overset{\bullet}{w}(\mathbf{x}) = \frac{1}{4\pi^2}\widehat{C}(\mathbf{0})\int_{-\infty}^{\infty}\widehat{A}(-\mathbf{m})\widehat{w}(\mathbf{m})d\mathbf{m} + \frac{1}{4\pi^2}\int_{\mathbf{n}\neq\mathbf{0}}\widehat{C}(\mathbf{n})\int_{-\infty}^{\infty}\widehat{A}(\mathbf{n}-\mathbf{m})\widehat{w}(\mathbf{m})e^{i\mathbf{n}\mathbf{x}}d\mathbf{m}d\mathbf{n}.$$

The first term in the right hand side is some constant independent of $\mathbf{x}$. If $\widehat{C}(\mathbf{n})$ is uni-modal then this term dominates. Therefore, it is possible to define k_1 to suppress the effect of the constant term in computer simulations.

Next we consider the role of k_2. Here we suppose that k_2 is the following function $k_2(\mathbf{x})$.

$$k_2(x, y) = \begin{cases} k_2 & |\mathbf{x}| \leq R, \\ 0 & \text{otherwise}, \end{cases}$$

that is, k_2 is effective only within the arbor radius. FT of $k_2(\mathbf{n})$ can be computed as follows.

$$\widehat{k_2}(\mathbf{n}) = \int_{-\infty}^{\infty} k_2(\mathbf{x})e^{-i\mathbf{n}\mathbf{x}}d\mathbf{x} = 2\pi k_2 R\frac{J_1(|\mathbf{n}|R)}{|\mathbf{n}|}.$$

This function takes the maximum value $2\pi k_2 R$ at the center and decreases to 0 toward to points $\mathbf{n}$ where $|\mathbf{n}| \approx 3.8/R \approx \pi/R$. After that this function oscillates and finally converges to 0. Thus, we could regard $\widehat{k_2}(\mathbf{n})$ as a sharp uni-modal function and that is "effective" only within the range $|\mathbf{n}| \leq \sigma$ (Fig. 8 left). Hence even if the $\widehat{C}(\mathbf{n})$ is uni-modal, by taking k_2 appropriately, we can make $(\widehat{C + k_2})(\mathbf{n})$ bi-modal (Fig. 8 right). Moreover, the peak of $(\widehat{C + k_2})(\mathbf{n})$ appears around $|\mathbf{n}| \approx \sigma$.

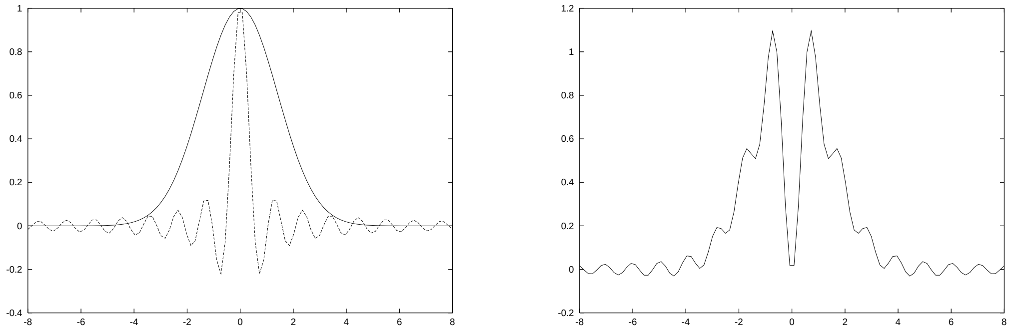

Figure 8: Left: The shape of $\widehat{C}(\mathbf{n})$ (real line) and that of $\widehat{k_2}(\mathbf{n})$ (dashed line). We assume that $\widehat{C}(\mathbf{n})$ is uni-modal. Right: the shape of $(\widehat{C - k_2})(\mathbf{n})$ for an appropriate k_2. As can be seen, the shape is bi-modal.

6 Conclusion

We considered Linsker's neural network model and investigated which features in this model are crucial to develop oriented receptive fields. We assumed that the correlation function is a Mexican-hat shape function while we introduced the arbor function explicitly and discussed its role in the development.

We computed FT of the modification rule to convert the convolution into a simple multiplication. Then we made the following observations. (1) FT of the correlation function extracts a set of Fourier coefficients of a certain frequency, and those coefficients are distributed on the circle of radius σ, (2) due to the symmetry of signs of Fourier coefficients, generative Fourier components are of which those frequencies are multiples of $\pi/2\sigma$, (3) among those generative components, the arbor function extracts only some of those and others degenerate, and (4) due to the arbor function, some sets of remaining Fourier components dominate the others, and these Fourier components eventually determine the shape of the receptive field. We further derived some condition under which an orientation receptive field is robustly developed.

These observations were also justified by our computer simulations, and finally we discussed the role of Linsker's parameters k_1 and k_2.

Acknowledgments

The author would like to thank Professor Osamu Watanabe at Tokyo Institute of Technology for supervising this work and for many helpful discussions. He also thank Dr. Shigeru Tanaka at the Institute of Physical and Chemical Research (RIKEN) and all the members in his laboratory for their valuable discussions and suggestions. The author is sincerely grateful to all referees and editors for their constructive criticism that helped greatly for improving this paper.

References

Feng, J., Pan, H., Roychowdhury, V. (1997). Linsker-type Hebbian Learning: A Qualitative Analysis on the Parameter Space. *Neural Network*, 10, 705–720.

Kammen, D., Yuille, A. (1988). Spontaneous Symmetry-Breaking Energy Functions and the Emergence of Orientation Selective Cortical Cells. *Biological Cybernetics*, 59, 23-31.

Linsker, R. (1986). From basic network principles to neural architecture (series). *Proc. Natl. Acad. Sci. USA*, 83, 7508–12, 8390–4, 8779–83.

Linsker, R. (1990). Self-Organization in a Perceptual System: How Network Models and Information Theory May Shed Light on Neural Organization. In S. Hanson and C. Olson, eds., *Connectionist Modeling and Brain Function*, MIT Press, 351-392.

MacKay, D., Miller, K. (1990). Analysis of Linsker's application of Hebbian rules to linear networks. *Network*, 1, 257–297.

Matsumoto, M., Nishimura, T. (1998). Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Trans. on Modeling and Computer Simulation*, 8, 3-30. (The source code can be retrieved from `http://www.math.keio.ac.jp/~matumoto/emt.html`)

Miller, K. (1994). 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.

Miyashita, M., Tanaka, S. (1992). A Mathematical Model for the Self-Organization of Orientation Columns in Visual Cortex. *NeuroReport*, 3 , 69–72.

Swindale, N. (1996). The development of topography in the visual cortex: a review of models. *Network*, 7 , 161–247.

Tang, D. (1989). Information-theoretic solutions to early visual information processing: Analytic results. *Physical Review A*, 40 , 6626–6635.