On the Stationary State of Kohonen’s Self-Organizing Sensory Mapping

H. Ritter and K. Schulten
Department of Physics, Technical University of Munich, D-8046 Garching, Federal Republic of Germany

Abstract. The stationary state of the self-organizing sensory mapping of Kohonen is investigated. For this purpose the equation for the stationary state is derived for the case of one-dimensional and two-dimensional mappings. The equation can be solved for special cases, including the general one-dimensional case, to yield an explicit expression for the local magnification factor of the map.

1 Introduction

Self-organizing sensory mappings play a crucial role in the development and maintenance of many functions of the nervous system and especially the brain. Different sensory inputs, such as tactile (Kaas 1983; Merzenich 1983), visual (Whitteridge 1973) and acoustic (Suga 1979; Pickles 1982) inputs, are known to be mapped onto different areas of the cerebral cortex in an orderly, topography-preserving fashion, i.e., similar inputs are mapped onto neighboring places in the cortex. These mappings are not genetically specified in a detailed manner but instead self-organize during the early stages of the formation of the nervous system. To some extent the mappings can remain plastic even later and adapt to subsequent changes in the environment or the sensors themselves. The degree of plasticity varies for different cortical mappings. For instance, the mapping from retina to cortex after its formation remains plastic for a relatively short period of time, whereas for the somatosensory map considerable plasticity has been found even in adult cats (Kaas 1983; Edelman 1983). In addition, different types of reorganisation after partial damage to afferent inputs have been observed (Kaas 1983).

Several algorithms for the formation of such mappings have been suggested (Edelman 1985; Takeuchi 1979; Willshaw 1976, 1979). In the following we will consider a proposal due to Kohonen (Kohonen 1982a, b). This proposal is not meant to model biological details but rather tries to capture the most essential features of such mappings for the benefit of remaining computationally tractable. The formation of the map is driven by a random sequence of sensory input signals whose probability distribution imprints on the final map in such a way that regions of the input signal space corresponding to frequent signal occurrences are mapped onto larger areas of the map than corresponding to rare input signals. Therefore the map acquires more important sensory regions at the expense of less important sensory regions.

Below we shall illustrate the algorithm, obtain an equation for the final (stationary) map in terms of the signal probability distribution and deduce the local magnification factor for special cases, including the general one-dimensional case.

2 The Model

As in (Kohonen 1982a, b) we consider a map \( \phi: A \mapsto B \) where \( A \) represents a lattice of neuronal units labelled by \( r \) and \( A \) is a spatially continuous sensory source with elements \( v \). A may represent, for example, the coordinate set of somatosensory receptors distributed densely over the body surface and \( B \) the set of those neuronal units of a layer in the cerebral cortex to which the somatosensory receptors are linked. The lattice \( B \) receives a sequence of input signals drawn randomly from \( A \), where \( v(r) \) is an input signal at location \( r \). Each input signal \( v(r) \) is received by all elements of \( B \) simultaneously. To each unit belongs a vector \( w(r) \) which determines the response of each unit upon arrival of a signal \( v(r) \). The response shall be given by 

\[ f(\sigma(v(r) - w(r))) \]

for \( \sigma \) a smooth real function peaked at \( x = 0 \) and of Gaussian type. Calling the union of all those points \( r \) which are closer to \( v(r) \) than to any other \( w(s,t) \), \( s \neq r \), the "receptive field" \( A \), of unit \( r \) (Fig. 1), we always have the maximal response at that unit \( r \), for which \( v(r) \) is in the "receptive field". The mapping \( \phi: A \mapsto B \) we are seeking is then specified as follows: the image of a vector \( v(r) \) is the particular unit \( w(r) \) which maximally responds to the signal \( v(r) \).

Initially the vectors \( w(r) \) are randomly distributed and therefore the receptive field of the individual units \( w(r) \) is distributed arbitrarily (e.g. randomly) in the input space \( A \). Each incoming signal \( v(r) \) corresponds arbitrarily to an input space \( v(r) = 1 \), \( 2 \), ... causes the following adaptation step to take place:

1) Selection of the unit \( r \) with maximal response upon \( v(r) \)

2) Modification of the receptive fields of unit \( r \) and all neighbouring units \( s \) according to

\[ w(s,t) = w(s,t) + h(s,t) \cdot \cdot (v(r) - w(s,t)) \]

For each \( r \), \( h(s,t) \) is peaked at \( s = r \) and of Gaussian type (either in each component if \( B \) is a high dimensional lattice or in the modulus of \( x \), whose width \( d(t) \) is a slowly decreasing function of \( t \). All units \( s \) at a distance to unit \( r \) exceeding \( d(t) \) receive only very little modification through \( l(t) \), whereas all closer units are modified notably so that their response to signal \( v(t) \) is changed in the spirit of Edelman's group selection theory (Edelman 1985) a unit might be interpreted as a
group of neurons. At each input the most responsive

As an illustration of this algorithm we show the

As a square array of $30 \times 30$ units and the

between those pairs $w(r_1, t)$ and $w(r_2, t)$ for which $r_1$

The function $h(t)$ chosen in our simulation was a

slowly decaying amplitude $a(t)$ times a Gaussian with

initial width $d(t)$ of 5 lattice spacings, slowly decreasing
to a final value of 2 lattice spacings after 5000 iterations

and then remaining there for the rest of the simulation.
The initial value of $a(t)$ was 0.5, exponentially decaying
to 0.1 during the first 5000 iterations and constant
adapt in the course of 50000 further iterations to this removal. Figure 8a shows the final distribution of the values \( w(x, t) \) over the input space and Fig. 8b depicts the array B with its units marked by the location of their center of maximal sensitivity. The "cortical region" which in Fig. 7b immediately after the "amputation" is seen to be deprived of inputs has now been "invaded" by sensory input mainly from the adjacent regions L, R and T respectively, whereas more distant parts have changed only slightly. This plasticity is very similar to that found for the somatosensory map in the experiment referred to above. The rearrangement of the map is accompanied by an increase in the map's local magnification factor for the adjacent parts of regions L, R and T, which results in a higher spatial sensory resolution there. This is also discernible from Fig. 8a, where an increase in the local density of the mesh-points \( w(x, t) \) in the surround of the "amputation" can be seen. This latter effect is also in good qualitative agreement with experimental observations (Merzenich 1983).

3 Equation for the Final (Stationary) Mapping

As is shown in Kohonen (1982a, c), repeating the above steps 1) and 2) and decreasing \( \text{d}(t) \) sufficiently slowly yields an ordered mapping from A onto the array of units such that neighboring units are sensitive to neighboring regions of A, irrespective of the initial values \( w(x, 0) \). The important dependence of the final mapping upon the probability distribution of the input signal \( v(t) \) was discussed only qualitatively in (Kohonen 1982a) and shall be supplemented here by a more quantitative treatment.

As long as \( d(t) \) is nonzero, \( w(x, t) \) undergoes a usually nonzero change at each time step. Given a configuration \( w(x, t) \) at time \( t \), the expectation value of its change up to time \( t+1 \) is

\[
\langle w(x, t+1) - w(x, t) \rangle = \langle \text{d}w(x, t) \rangle
\]

(1)

where \( \langle \ldots \rangle \) denotes the average over all possible values of \( v(t) \) and \( h(t, d(t)) \) stands for the former \( h(t, r) \) to make the \( d \)-dependence explicit.

Keeping \( d(t) = 1 \) fixed for the moment, we shall call a configuration \( w(x, t) \) an equilibrium configuration, if \( w(x, t) = w(x, 0) \) yields a vanishing expectation value in (1). We want to consider the equilibrium configuration in the limit of vanishing fluctuations, i.e.

\[
w(x, t) = \lim_{\text{D} \to 0} w(x, t).
\]

The following analysis will be restricted to the case of A and B being of the same dimension \( n \) (although the algorithm is capable of establishing a mapping between different dimensional A and B either, see (Kohonen 1984) and the validity of two main assumptions:

i) We assume that for sufficiently many units and all sufficiently small \( d \) the equilibrium configurations \( w(x, t) \) are sufficiently slowly varying with \( t \) to allow replacing them by corresponding smooth functions over a continuum of \( r \)-values and consequently setting \( \mu^{-1} = \text{w}(\mu) \). This basically assumes that the topological ordering of the final stage has already occurred.

ii) We will assume bijective equilibrium configurations \( w(x, t) \). This is a reasonable assumption, since the discrete algorithm has the tendency to avoid mapping the same subregion of the signal space A to different parts of B.

In addition we require \( h(x, d) \) to be Gaussian of type with width of order \( d \) and with vanishing first and isotropic second moments for all \( a, d, i, e \).

\[
\langle h(x, d) \rangle = \frac{\partial Q}{\partial D} + \frac{\partial d}{\partial D}
\]

(7)

\[
J \cdot \text{V} \ln(Q \cdot D) = -J \cdot d \cdot w(x, t)
\]

(8)

We are now going to derive a necessary and sufficient differential equation for \( \text{w}(x) \) to be stationary. We start with the equilibrium condition for \( \text{w}(x) \).

\[
\langle h(x, d) \rangle = \frac{\partial Q}{\partial D} + \frac{\partial d}{\partial D} = 0
\]

(9)

or, introducing the Jacobian \( J(x) = \partial Q / \partial d \).

\[
J \cdot \text{V} \ln(Q \cdot D) = -J \cdot d \cdot \text{w}(x, t)
\]

(10)

As we are only interested in the limit \( d \to 0 \), we shall henceforth denote \( \text{w}(x, t) \) by \( \text{w}(x) \) solely. An alternative form of (8) is obtained via

\[
\ln(Q \cdot D^2) = -J \cdot d \cdot \text{w}(x, t) + \frac{1}{2} \ln(D - J \cdot d \cdot \text{w}(x, t))
\]

(9a)

or

\[
\ln(Q \cdot D^2) = -\frac{1}{2} \cdot \text{w} \cdot \text{swn}(d J)\]

(10a)

where \( u \) is given by

\[
u = \text{det}(u) \cdot J^{-1} \cdot d \cdot \text{w}(x, t)\]

(11a)

In two dimensions with \( u = (a, b) \), this can be written more symmetrically as

\[
u = (a, b) \cdot \text{swn}(d J)
\]

(12a)

Equations (8) or (9), together with suitable boundary conditions, determine the equilibrium configuration \( w(x) \), which in turn represents the inverse of the original map \( A \to B \), since \( w(x) \) is the center of A of maximal sensitivity of unit \( r \rightarrow B \).

4 Discussion

Although the nonlinearity of (8) (9) makes a general discussion unfeasible, in one and two dimensions some consequences concerning the relationship between the local magnification factor and the driving probability distribution \( P \) may be derived immediately.

\[
\text{w}(x) \text{ represents the inverse of the map } A \to B \text{, the local magnification factor } M \text{ of the latter is given by}
\]

\[
M(x) = \frac{\partial Q}{\partial D} + \frac{\partial d}{\partial D}
\]

(13a)

\[
J \cdot \text{V} \ln(Q \cdot D) = -J \cdot d \cdot \text{w}(x, t)
\]

(14a)
$M = 1/D$ (cf. (4)). It has a simple dependence on the density $P(w(t))$ of inputs in at least two cases. The first case arises for $w$ such that $v$ vanishes. Then $Q_D = P(w)$ const. and, therefore, (employing the identity $P(w(t)) = Q(t))$

$$M(w) = D^{-1} Q(w) \alpha,$$  

where $w$ vanishes whenever $A$, $B$ are either (i) both one-dimensional or (ii) of rectangular shape and $P$ is a product $P(w) = P_A(a) P_B(b)$ with $w = (a, b)^T$. In the latter case the choice $a = a(x), b = b(y)$ splits (9) into two first order equations with $x$ and $y$ decoupled, yielding

$$x = c_1 \int A(x) dx$$

and

$$y = c_2 \int B(y) dy.$$  

The four integration constants $c_1, c_2, a_0, b_0$ are fixed by a particular choice for the (arbitrary) starting point and the (arbitrary) scale for the labelling of the units in the $x$- and $y$-directions, respectively.

The second case in which a relationship between $P(w)$ and $w$ can be established is when $w$ can be represented by a complex function

$$w = (Re w, Im w)^T$$

with $w$ analytic in $z = x + iy$. This yields $QD = D^{-1} P(w)$ and therefore

$$M(w) = \alpha P(w).$$

An example is given by $P(w) = w^2/(\|w\|^2)$ and

$$A = \{w e^{-w} \|w\|^2 < 1 \& w > 0\},$$

$$B = \{0, N\} \times \{0, N\}.$$  

For this choice follows for $a(w)$ defined through (14)

$$a(z) = \exp(z - 1 - z).$$

This yields a map $\phi$ from the semi-annulus $A$ onto the square $B$. Such a kind of map connects, for example, the retina with the visual cortex.

In general, in the case of a two-dimensional mapping the magnitude factor $M(w)$ of the stationary map is not expressible as a simple function of the local probability density $P(w)$ of the driving input as is implied in (Kohonen 1982c, 1984). Only in the one-dimensional case such a relationship can be derived. The derivation yields $M(w) \propto P(w)^{1/2}$, a result which may be in contrast to the intuitive, but incorrect expectation $M(w) \propto P(w)$ suggested in (Kohonen 1984).

To test our findings we simulated Kohonen's map for the case of a one-dimensional lattice $B$ of 1000 units and an interval $A = [0, 1]$. The probability density of inputs from $A$ was chosen linearly, i.e. $P(w) = 2w$. Figure 9 represents the result of this simulation. Initially we chose the map $w(r) = \sqrt{r}$, whose magnification factor is proportional to $P$. After 100 000 iterations the map has developed away from its initial configuration and reached the equilibrium curve $w(r) = r^{1/3}$ corresponding to a magnification factor $M(w) \propto P(w)^{1/2}$ as predicted by (11).

5 Conclusion

We have derived an equation for the equilibrium state of a self-organizing topographic mapping due to Kohonen and for some special cases derived analytical expressions of the local magnification factor in terms of the probability density of the driving input. It is shown, that the local magnification factor in the one-dimensional case is proportional to $P^{1/2}$, whereas in two dimensions no general local expression in terms of the probability density can be given.

References


