Topology Preservation in Self-Organizing Feature Maps : General Definition and Efficient Measurement

Th. Villmann[†] R. Der[†] M. Herrmann[†] Th. Martinetz[‡]
[†]Universität Leipzig, Inst. für Informatik
D–04109 Leipzig, Augustusplatz 10/11, BR Deutschland
[‡]Siemens AG, ZFE
D–81730 München, BR Deutschland

to appear in *Lecture Notes in Computer Science*

Abstract

In this paper we present a new approach to the problem of measuring the topology preservation in SOFM. We introduce a precise definition of the meaning of that property and derive the so-called topographic function for its measuring based on the receptive fields of the neural units using explicitly the structure of the given data manifold.

1 Introduction

The capability of topology–preserving mapping of a data manifold onto a lattice of neural units is one of the advantages of Kohonen's self-organizing feature map (SOFM) [8],[9], [12]. This property can be used in a variety of information processing tasks, ranging from classification over robotics to data reduction and knowledge processing. To each neural unit a reference or synaptic weight vector is assigned, defining the receptive field consisting of all data points which are matched best by this reference vector.

Various qualitative and quantitative methods for characterizing the degree of topology preservation [1], [5], [16] have been proposed. However, all these approaches use only an intuitive definition of topology preservation based on the consideration of the weights of the neural units. These approaches can not distinguish a correct folding due to the folded data manifold from a folding due to a topological mismatch between data manifold and neural lattice. The problem is shown in Fig.1. In both the linear and nonlinear case of M the situations of the weight vector s of the neural units are the same and, hence, the methods based on the consideration of the weight vectors would indicate a dimensional conflict in both cases. However in the nonlinear case the map has been formed correctly. Particularly, when using the SOFM for non-linear principle component analysis one has to have a means to distinguish between these two cases.

In the case of the input and network space both being one-dimensional the definition of topology preservation is trivial: there are essentially two ordered arrangements of neurons, one with increasing, the other with decreasing neural indices when moving through the input space. For higher dimensionalities it is intuitively clear what topology preservation should mean, although no formal definition has been given. By the use of the formalism of the mathematical topology we derived a general definition of this property and a new approach for quantifying it using explicitly the structure of the data manifold.

Kohonen's algorithm determines a SOFM describing the map $\Psi_{M\longrightarrow A}$ from a data manifold $M \subseteq \Re^d$ onto a d_A-dimensional lattice $A \subseteq \Re^{d_A}$ of neural units and the inverse mapping $\Psi_{A\longrightarrow M}$. The structure of the lattice is defined by its connectivity graph $\mathbf{C}^A$. The map $\mathcal{M}_A = (\Psi_{A\longrightarrow M}, \Psi_{M\longrightarrow A})$ of M formed by A is then determined by

$$\mathcal{M}_A = \begin{cases} \Psi_{M\to A} \ : \ M \longrightarrow A \quad v \in M \longmapsto i^*(v) \in A \\ \\ \Psi_{A\to M} \ : \ A \longrightarrow M \quad i \in A \longmapsto w_i \in M \end{cases} \tag{1}$$

with $i^*(v)$ as the neural unit with its synaptic weight vector $w_{i^*(v)}$ closest to v, i.e., with

$$\left\| w_{i^*(v)} - v \right\| \leq \|w_j - v\| \quad \forall j \in A \tag{2}$$

Kohonen's self-organizing feature map algorithm distributes the synaptic weight vectors w_i such, that the map $\mathcal{M}_A$ of M formed by A is as topology preserving as possible. The reference vectors w_i are adapted in a learning step according to

$$\triangle\, w_i = \epsilon h_{i^*,i}\,(v - w_i) \quad \forall i \in A, \tag{3}$$

where $v \in M$ is the presented stimulus vector, $i^*(v)$ is defined again by eq. (2) and the neighborhood function

$$h_{i^*,i} = \exp\left(-\frac{\|i^* - i\|_A^2}{2\sigma^2}\right) \tag{4}$$

determines the neighborhood range in A by the choice of the radius σ. $\|\cdot\|_A$ denotes the Euclidean distance in A. ϵ is the learning parameter.

2 Definition of Topology Preservation in SOFM

We want to call a map $\mathcal{M}_A = (\Psi_{A\to M}, \Psi_{M\to A})$ of M "topology preserving", if both the mapping $\Psi_{M\to A}$ from M to A as well as the inverse mapping $\Psi_{A\to M}$ from A to M is neighborhood preserving. Hence, to determine whether a SOFM is topology preserving we have to measure these two neighborhood preservations. Per definition we regard the mapping $\Psi_{M\to A}$ from M to A as being neighborhood preserving if reference vectors w_i, w_j which are adjacent on M, belong to vertices i, j, which are neighbors in A. On the other hand, the inverse mapping $\Psi_{A\to M}$ from A to M is neighborhood preserving if adjacent vertices i, j are mapped onto locations w_i, w_j which are neighbors on M. How can we define *neighborhood of neural units i, j in A* and *neighborhood of reference vectors w_i, w_j on M* in a way that the intuitive understanding of topology preservation of a SOFM is captured?

The basic idea is to describe the property of topology preservation in terms of the mathematical topology [14]. Then the property of topology preservation of a map may formulate by the continuity of this map between topological spaces. At first we introduce some helpful concepts. The induced Voronoi diagram $\mathcal{V}_M$ of a subset $M \subseteq \Re^d$ and its dual the Delaunay graph (Voronoi graph) $\mathcal{D}_M$ with respect to a set $S = \{w_1, \ldots, w_N\}$ of points $w_i \in M \subseteq \Re^d$ is given by the masked Voronoi polyhedra

$$\tilde{V}_i = \{x \in M \mid \|x - w_i\| \leq \|x - w_j\| \quad j = 1 \ldots N, \quad j \neq i\} \tag{5}$$

as shown in [10], [11]. We remark that the Voronoi polyhedra are closed sets. The cells form a complete partitioning of M in the sense that $M = \cup_{i=1}^{N} \tilde{V}_i$. The induced Voronoi diagram $\mathcal{V}_M$ uniquely corresponds to its Delaunay graph $\mathcal{D}_M$ [4]. Two Voronoi cells $\tilde{V}_i$, $\tilde{V}_j$ are connected in $\mathcal{D}_M$ if and only if the intersection of it is non-vanishing, i.e. $\tilde{V}_i \cap \tilde{V}_j \neq \emptyset$ [4], [15]. This allows us to define a graph metric in $\mathcal{D}_M$ as the minimal path length in the graph. In the general case the Voronoi diagram $\mathcal{V}$ of $\Re^d$ with respect to S is given by the Voronoi cells defined by

$$V_i = \left\{x \in \Re^d \mid \|x - w_i\| \leq \|x - w_j\| \quad j = 1 \ldots N, \quad j \neq i\right\} \tag{6}$$

Using the above introduced concepts we are now able to define in a general manner what topology preservation for arbitrary lattices A with the connectivity graph $\mathbf{C}^A$ means. However, a proper definition of the topology preservation of the two maps of $\mathcal{M}_A$ which allows small distortions as depicted in Fig.2 requires two different kinds of topology in the lattice A:

Definition 2.1 *Suppose A to be a network of N neurons which are situated at points $i = (i_1, \ldots, i_{d_A}) \in \Re^{d_A}$ with reference or synaptic weight vectors $w_i \in M \subseteq \Re^d$. The connectivity graph $\mathbf{C}^A$ of A defines the structure of A.*

Consider for the moment A to be a set of points in $\Re^{d_A}$. A (discrete) topology $\mathcal{T}_A^+$ in the set A is induced by the graph metric in $\mathbf{C}^A$. $\mathcal{T}_A^+$ is said to be the **strong neighborhood topology in** *A, and $\left(A, \mathcal{T}_A^+\right)$ is a topological space.*

Definition 2.2 *Let $\mathcal{V}$ be the Voronoi diagram of $\Re^{d_A}$ with respect to A and $\mathcal{D}_A$ be its dual Delaunay graph. $\mathcal{D}_A$ is equipped with the graph metric that in turn induces the (discrete) topology $\mathcal{T}_A^-$ in $\mathcal{V}$ and, hence, also in A. $\mathcal{T}_A^-$ is said to be the* **weak neighborhood topology in** *A, and $\left(A, \mathcal{T}_A^-\right)$ is a further topological space defined on the set A.*

Remark 2.1 *In the case of a rectangular lattice the weak neighborhood topology $\mathcal{T}_A^-$ is weaker then the strong neighborhood topology $\mathcal{T}_A^+$ also in the sense of the mathematical topology [6].*

In the next step we introduce a topology in the set of the synaptic weight vectors on the basis of their receptive fields, which again allows us to describe the neighborhood relationships between two vectors.

Definition 2.3 *Let $\Psi_{A \to M} : A \longrightarrow M^A \subset M \subseteq \Re^d$ with $M^A = \{w_i,\ i \in A\}$ be a map attributing to each neuron i a specific vector $w_i \in M^A$. Furthermore, let $\mathcal{V}_M$ be the induced Voronoi diagram of M with respect to $M^A = \{w_i,\ i \in A\}$. Let $\mathcal{G}_M$ be the dual Delaunay graph of $\mathcal{V}_M$. $\mathcal{G}_M$ is equipped with the graph metric that in turn induces the (discrete) topology $\mathcal{T}_{M^A}$ in $\mathcal{G}_M$ and, hence, also in M^A. $\mathcal{T}_{M^A}$ is said to be the $\Psi_{A \to M}$*–**induced neighborhood topology** *in M^A and $\left(M^A, \mathcal{T}_{M^A}\right)$ is a topological space.*

Now the topology preservation of a map can be expressed by the following definition:

Definition 2.4 *The map $\mathcal{M}_A = (\Psi_{A \to M}, \Psi_{M \to A})$ with $\Psi_{M \to A} : \Re^d \supseteq M \longrightarrow A$ defined by $i(v) = \arg\left(\min_{j \in A} \|v - w_j\|\right)$ is said to be* **topology preserving** *if both $\Psi_{M \to A} : \left(M^A, \mathcal{T}_{M^A}\right) \longrightarrow \left(A, \mathcal{T}_A^-\right)$ and $\Psi_{A \to M} : \left(A, \mathcal{T}_A^+\right) \longrightarrow \left(M^A, \mathcal{T}_{M^A}\right)$ are continuous maps of the respective topological spaces. Irrespective of the different topologies $\Psi_{M \to A}$ is the inverse mapping of $\Psi_{A \to M}$.*

We have immediately the following two corollaries for the most important cases of rectangular and hexagonal (triangular) lattices, respectively:

Corollary 2.1 *In the case of a rectangular d_A-dimensional lattice A of neurons the strong topology is induced by the Euclidean norm $\|\cdot\|_{A,E}$ in A or the summation-norm $\|\cdot\|_{A,\Sigma} = \sum_{j=1}^{d_A} \left|(\cdot)_j\right|$, and the weak topology is induced by the maximum-norm $\|\cdot\|_{A,\max} = \max_{j=1}^{d_A} \left|(\cdot)_j\right|$.*

Corollary 2.2 *In the special case of A being a hexagonal (triangular) lattice the weak and strong topology coincide. Hence, the definition of topology preservation relies on a single topology in the net [10] which corresponds to the strong neighborhood topology.*

For measuring the degree of topology preservation of $\mathcal{M}_A = (\Psi_{A\to M}, \Psi_{M\to A})$, we have to proof now the continuity of the maps $\Psi_{A\to M}$ and $\Psi_{M\to A}$ as defined above. Therefore based on the definitions 2.1, 2.2, 2.3 and 2.4 we introduce a measure, what we call the topographic function, which determines the degree of topology preservation of $\mathcal{M}_A = (\Psi_{A\to M}, \Psi_{M\to A})$. A first simpler version of the topographic function was proposed in [13].

For each unit i we define

$$\begin{aligned} f_i(k) &\stackrel{def}{=} \#\left\{j \mid \|i-j\|_{\mathcal{T}_A^-} > k \; ; \; \|w_i - w_j\|_{\mathcal{T}_{MA}} = 1\right\} \\ f_i(-k) &\stackrel{def}{=} \#\left\{j \mid \|i-j\|_{\mathcal{T}_A^+} = 1 \; ; \; \|w_i - w_j\|_{\mathcal{T}_{MA}} > k\right\} \end{aligned} \tag{7}$$

with $k = 1, \ldots, N-1$. $\#\{\cdot\}$ denotes the cardinality of the set. $\|\cdot\|_{\mathcal{T}}$ is a norm based on the topology $\mathcal{T}$. Looking at a neural unit i, $f_i(k)$ with $k > 0$ determines the continuity of $\Psi_{M\to A}$ and $f_i(k)$ with $k < 0$ determines the continuity of $\Psi_{A\to M}$ as defined above. The topographic function of the neural lattice A with respect to the input manifold M is then defined as

$$\Phi_A^M(k) \stackrel{def}{=} \begin{cases} \frac{1}{N}\sum_{j\in A} f_j(k) & ; \quad k > 0 \\ \Phi_A^M(1) + \Phi_A^M(-1) & ; \quad k = 0 \\ \frac{1}{N}\sum_{j\in A} f_j(k) & ; \quad k < 0 \end{cases} \tag{8}$$

and we remark:

Remark 2.2 *We obtain $\Phi_A^M \equiv 0$ and, particularly, $\Phi_A^M(0) = 0$ if and only if the SOFM is perfectly topology preserving. The largest $k^+ > 0$ for which $\Phi_A^M(k^+) \neq 0$ holds yields the range of the largest fold if the effective dimension of the data manifold M is larger than the dimension d_A of the lattice A. The smallest $k^- < 0$ for which $\Phi_A^M(k) \neq 0$ holds yields the range of the largest fold if the effective dimension of the data manifold M is smaller then the dimension d_A of the lattice A. Small values of k^+ and k^- indicate that there are only local conflicts, large values indicate a global character of the dimensional conflict.*

Fig.3 shows a map of a squared data manifold onto a chain of 100 neural units with their receptive fields. The folds are involved all over the whole chain and, hence, the topographic function vanishes only for k-values greater then $k^+ = 98$.

In Ref. [7] a definition was presented of what it means for a topographic map to be *ordered*, which finally should be discussed in the light of the definition studied in the present paper. The following only one-dimensional definition there was introduced

Definition 2.5 *We consider a chain of N neural units i with weight vectors w_i and receptive fields $R_i = \tilde{V}_i$ as defined in 5. Let $\eta_i(v)$ be a activity function of the ith neuron with respect to a stimuli vector $v \in M \subseteq \Re^d$, for instance the negative distance $-\|w_i - v\|$ or the inverse distance $\left(\frac{1}{\|w_i - v\|}\right)$. Let $X = \left\{x_j \in M \subseteq \Re^d \mid j = 1, \ldots, n\right\}$ be a set of points, such that $x_1 \stackrel{Rel}{\circ} x_2 \stackrel{Rel}{\circ} x_3 \stackrel{Rel}{\circ} \ldots \stackrel{Rel}{\circ} x_n$ and for all neural units i exists a $x_j \in X$ for which $x_j \in R_i$ holds. There $\stackrel{Rel}{\circ}$ is an arbitrary suitably chosen relation, not necessarily transitive. The system is said to implement a* one–dimensional ordered mapping *if for $i_1 > i_2 > i_3 > \ldots$*

$$\begin{array}{rcl} \eta_{i_1}(x_1) & = & \max_{i=1,\ldots,N} \eta_i(x_1) \\ \eta_{i_2}(x_2) & = & \max_{i=1,\ldots,N} \eta_i(x_2) \\ \eta_{i_3}(x_3) & = & \max_{i=1,\ldots,N} \eta_i(x_3) \\ & \vdots & \\ & etc. & \end{array} \tag{9}$$

holds.

We have given in this paper an explicit order relation $\stackrel{Rel}{\circ}$ based on an underlying topology, in contrast to the requirement of the existence of such a relation, which was formulated for the one–dimensional case only. The proposed straight–forward generalization of the definition (9) to higher dimensions is (as pointed out in [3]) by no means trivial. In fact, one need to make the relation $\stackrel{Rel}{\circ}$ more explicit in order to find whether the definition (9) applicable for higher–dimension, too. In this sense we gave the justification of Kohonen's early and not yet otherwisely worked out set–up. A further issue is the topology preservation of the map $\Psi_{A \to M}$ which has been pointed out to be as well crucial for strict definition of the intuitive notion of topology preservation.

3 Computing the Topographic Function Φ_A^M

Computing Φ_A^M requires to determine the topological relationship in M^A, i.e. to determine whether two receptive fields R_i, R_j are adjacent on the given manifold M. A way to determine the adjacency of two receptive fields $R_i = \tilde{V}_i$, $R_j = \tilde{V}_j$ with $\tilde{V}_i$, $\tilde{V}_j$ as defined in (5) has been proposed in [10]. Let $\mathbf{C}$ be a connectivity matrix determining connections between units $i, j \in A$ (in addition to the connectivity matrix defined by the fixed lattice structure). Initially, the elements $\mathbf{C}_{ij} \in \{0, 1\}$ of $\mathbf{C}$ are set to zero. Simply by sequentially presenting input vectors $v \in M$ and each time connecting (setting $\mathbf{C}_{i^*j^*} = 1$) those two units i^*, j^*, the reference vectors w_{i^*} and w_{j^*} of which are closest and second closest to v, leads to a connectivity matrix $\mathbf{C}_{ij}$ for which

$$\lim_{t\to\infty} \mathbf{C}_{ij} = 1 \quad \Leftrightarrow \quad R_i \cap R_j \neq \emptyset \tag{10}$$

is valid. This algorithm is based on the competitive "Hebbian rule" [10]. It can be shown that the resulting connectivity structure connects units and only units the receptive fields of which are adjacent [11]. Then the structure of the topology can easily obtained as distance matrix $\mathbf{D}$ with $\mathbf{D}_{ij} = \|w_i - w_j\|_{T_{MA}}$ by determining the minimal ways in the dual graph using the connectivity matrix $\mathbf{C}$ [2]. This allows to rewrite eq. (7) into

$$\begin{aligned} f_i(k) &= \#\left\{ j \mid \|i - j\|_{T_A^-} > k \,; \mathbf{D}_{ij} = 1 \right\} \\ f_i(-k) &= \#\left\{ j \mid \|i - j\|_{T_A^+} = 1 \,; \mathbf{D}_{ij} > k \right\} \end{aligned} \tag{11}$$

for k-values in the range of $k = 1, \ldots, N_{\max}$. After a SOFM has been formed, we then can determine Φ_A^M by the following algorithm:

1. present an input vector $v \in M$
2. determine the two nearest reference vectors w_{i^*}, w_{j^*}.
3. connect the units i^*, j^*, i.e., set $\mathbf{C}_{i^*j^*} := 1$ and go to step 1

After a sufficient number of presented input vectors v the algorithm yields a connectivity matrix $\mathbf{C}$ for which eq. (10) is valid. Then, the matrix $\mathbf{C}$ can be used to calculate $\mathbf{D}$ and subsequently the topographic function Φ_A^M according to eq. (11) and eq. (8).

4 Conclusion

We presented a general definition of topology preservation in SOFM's and a novel approach to the problem of measuring the topology preservation. The approach is based on the neighborhood relations between receptive fields. The introduced topographic function is an improvement over the topographic product suggested in [1] since it determines the degree of topology preservation by considering explicitly the given input manifold M. Examples also in comparison to the topographic product will be discussed in [14].

Acknowledgement: THE REPORTED RESULTS ARE BASED ON WORK DONE IN THE PROJECT 'LADY' SPONSORED BY THE GERMAN BMFT UNDER GRANT 01 IN 106B/3.

References

[1] Bauer, H.-U.; Pawelzik, K.: Quantifying the Neighbourhood Preservation of Self-Organizing Feature Maps. *IEEE Trans: on Neural Networks* **3(4)**, 570-579, (1992).

[2] Biess, G.: Graphentheorie. BSB B. G. Teubner Verlagsgesellschaft Leipzig, (1988).

[3] Cottrell, M.; Fort, J. C.; Pages, G.: Two or Three Things That We Now About the Kohonen Algorithm. Proceedings of the European Symposium on Artificial Neural Networks (ESANN), Brussels (1994).

[4] Delaunay, B.: Sur la Spere Vide. *Bull. Acad. Sci. USSR (VII)*, Classe Sci. Mat. Nat., 793-800, (1934).

[5] Der, R.; Herrmann, M.; Villmann, Th.: Time Behavior of Topological Ordering in Self-Organized Feature Mapping. submitted to *Biol. Cybern.* (1993).

[6] Kantorowich, L. W.; Akilov, G. P.: Funktionalanalysis in normierten Räumen. Akademie-Verlag, Berlin, (1978).

[7] Kohonen, T.: A Simple Paradigm for the Self-organized Formation of Structured Feature Maps. In : Competition an Cooperation in Neural Nets. (*Lecture Notes in Biomathematics*, Eds.: S. Amari and M. A. Arbib) Springer-Verlag Berlin, Heidelberg, New York (1982).

[8] Kohonen, T.: Self-Organization and Associative Memory, Springer Series in Information Science 8, Springer, Berlin, Heidelberg (1984).

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

[10] Martinetz, Th.: Competitive Hebbian Learning Rule Forms Perfectly Topology Preserving Maps. Proc. of the International Conference on Artificial Neural Networks 1993, Eds. St. Gielen and B. Kappen, Springer, London, Berlin Heidelberg (1993).

[11] Martinetz, Th.; Schulten, K.: Topology Representing Networks. to appear in *Neural Networks* (1994).

[12] Ritter, H.; Martinetz, Th.; Schulten, K.: Neural Computation and Self–Organizing Maps. Addison Wesley: Reading, Mass., (1992).

[13] Villmann, Th.; Der, R.; Martinetz, Th.: A Novel Approach to Measure the Topology Preservation of Feature Maps. to appear in Proc. of the International Conference on Artificial Neural Networks 1994, Sorrento (1994).

[14] Villmann, Th.; Der, R.; Herrmann, M.; Martinetz, Th.: Measuring Topology Preservation in Self-Organizing Feature Maps. submitted to *IEEE Trans. on Neural Networks* (1994).

[15] Voronoi, G.: Nouvelles Aoolications des Parametres a la Theorie des Formes Quadratiques. Deuxieme Memorie: Recherches sur les Paralleloedres Primitifs, J. reine angew. Math., **134,** 198-287, (1908).

[16] Zrehen, St.: Analyzing Kohonen Maps With Geometry. Proc. of the International Conference on Artificial Neural Networks 1993, Eds. St. Gielen and B. Kappen, Springer, London, Berlin, Heidelberg (1993).

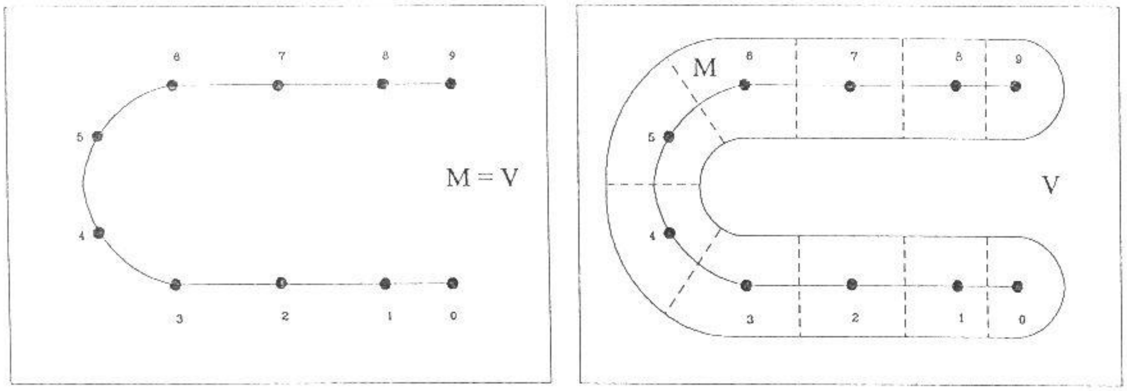

Figure 1: Example of a linear (left, $M = V$) and nonlinear (right, $M \subset V$) data manifold with the hypothetical positions of the images of the neural units

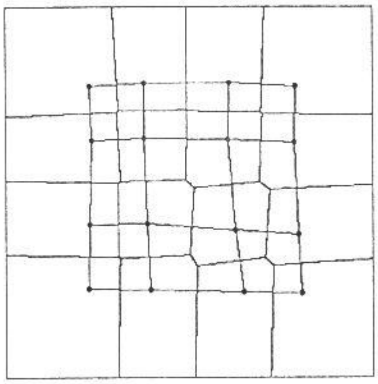

Figure 2: General case of the receptive fields of a squared lattice of neural units

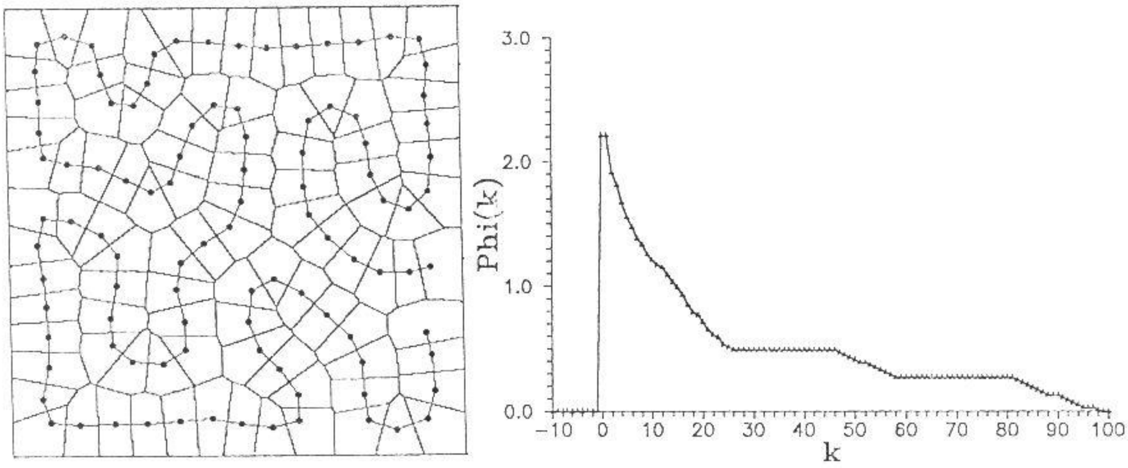

Figure 3: Plot of a map of a squared input space onto a chain of 100 neural units, the receptive fields of the units are shown (left); plot of the topograpcic function of the map (right)