In: Proceedings of the ICANN-93, Gielen S. and Kappen B. (eds.), Springer, Heidelberg, pp. 427-434

Competitive Hebbian Learning Rule Forms Perfectly Topology Preserving Maps

Thomas Martinetz

Siemens AG
Corporate Research and Development
Neural Networks Group
8000 München 83, Germany

Abstract: The problem of forming perfectly topology preserving maps of feature manifolds is studied. First, through introducing "masked Voronoi polyhedra" as a geometrical construct for determining neighborhood on manifolds, a rigorous definition of the term "topology preserving feature map" is given. Starting from this definition, it is shown that a network G of neural units i, $i = 1, ..., N$ has to have a lateral connectivity structure $\mathbf{A}$, $A_{ij} \in \{0, 1\}$, $i, j = 1, ..., N$ which corresponds to the "induced Delaunay triangulation" of the synaptic weight vectors $\mathbf{w}_i \in \Re^D$ in order to form a perfectly topology preserving map of a given manifold $M \subseteq \Re^D$ of features $\mathbf{v} \in M$. The lateral connections determine the neighborhood relations between the units in the network, which have to match the neighborhood relations of the features on the manifold. If all the weight vectors $\mathbf{w}_i$ are distributed over the given feature manifold M, and if this distribution resolves the shape of M, it can be shown that Hebbian learning with competition leads to lateral connections $i - j$ ($A_{ij} = 1$) that correspond to the edges of the "induced Delaunay triangulation" and, hence, leads to a network structure that forms a perfectly topology preserving map of M, independent of M's topology. This yields a means for constructing perfectly topology preserving maps of arbitrarily structured feature manifolds.

1. Introduction

Topology preserving feature maps play an important role in a variety of natural as well as artificial neural information processing systems [1-3]. By projecting input patterns onto a network of neural units such that similar patterns are projected onto adjacent units and, vice versa, such that adjacent units code similar patterns, a representation of the input patterns is achieved which in postprocessing stages allows one to exploit the similarity relations of the input patterns. Examples of topology preserving feature maps in the nervous system are the retinotopic map in the visual cortex [4], the mapping from the body surface onto the somatosensory cortex [5], or the tonotopic maps in the auditory cortex [6]. As components of artificial neural information processing systems topology preserving feature maps have been applied successfully in speech processing [7, 8], image processing [9], and robotics [10].

A number of neural network models for adaptively forming topology preserving feature maps have been proposed [11-14]. A model which provides a very compact procedure and, therefore, has found widespread application in artificial neural information processing systems is Kohonen's self-organizing feature map [2, 13]. This algorithm requires that one first chooses a graph (network) G, usually a one-, two-, or three-dimensional lattice; in a subsequent adaptation step, pointers (synaptic weight vectors) $\mathbf{w}_i$ which are assigned to the vertices (neural units) i of G are distributed over a given feature manifold $M \subseteq \Re^D$ in such a way, that (i) pointers lie on M, and (ii) pointers of vertices which are adjacent in G are assigned to locations which are close on M. To obtain a topology preserving map, it is necessary to choose a graph G, the topological structure of which matches the topological structure of the given feature manifold M.

In many applications, however, the feature manifold M is a submanifold of a high-dimensional space and may neither be known *a priori* nor topologically simple enough for prespecifying a correspondingly structured graph G. For these applications it would be highly desirable to have a procedure which adapts the topology of the graph G to the topology of the given manifold M. An approach to this problem has been introduced by Kohonen and coworkers [15]. They take the minimum spanning tree between the pointers $\mathbf{w}_i$ as the graph G. Another approach has been proposed by Fritzke [16]. His approach distributes two-dimensional, triangular cell structures over the manifold M for forming an appropriate graph G. In this paper we will describe an approach which is based on the so-called Delaunay triangulation [17] of the pointers $\mathbf{w}_i$ and employs a competitive version of the Hebb rule for forming the graph G. A preliminary version was presented in [18]. Starting from a rigorous definition of topology preservation, which is given in the next section, we show that the approach presented forms network structures G which preserve the topology of given feature manifolds completely.

2. A rigorous definition of topology preservation

Given a feature manifold M. Which graphical structure forms a perfectly topology preserving map of M? To answer this question, we first have to define exactly when topology preservation is given. The problem is that adjacency of vertices i in a graph G is clearly defined; however, a definition for adjacency of pointers $\mathbf{w}_i$ on M which is in agreement with our intuitive understanding of topology preservation is not obvious. This is why in previous contributions on topology preserving feature maps the interpretation of "topology preservation" has usually been left to the reader's intuition.

An exception is the trivial one-dimensional case. Obviously, two points $\mathbf{w}_i, \mathbf{w}_j \in \Re$ are neighboring if there is no point $\mathbf{w}_k$ in between. Expressed in terms of Voronoi polyhedra an equivalent definition is: two points $\mathbf{w}_i, \mathbf{w}_j \in \Re$ are neighboring if their Voronoi polyhedra V_i, V_j are adjacent, i.e., if $V_i \cap V_j \neq \emptyset$ with

$$V_i = \{\mathbf{v} \in \Re^D \mid \|\mathbf{v} - \mathbf{w}_i\| \leq \|\mathbf{v} - \mathbf{w}_j\| \; j = 1, \ldots, N\} \qquad i = 1, ..., N. \tag{1}$$

In these terms a generalization to higher dimensional embedding spaces $\Re^D$ is straightforward. However, since we need a definition of neighborhood of points *on a manifold* M, we first introduce the *masked Voronoi polyhedron*. The masked Voronoi polyhedron $V_i^{(M)}$ is the part of V_i which is also part of M, i.e., $V_i^{(M)} = V_i \cap M$. The superscript indicates the dependence of the masked Voronoi polyhedron on the given manifold M. By using the neighborhood of the masked Voronoi polyhedra $V_i^{(M)}$, $V_j^{(M)}$ instead of the neighborhood of the Voronoi polyhedra V_i, V_j for determining the neighborhood of the points $\mathbf{w}_i$, $\mathbf{w}_j$ on M, we ensure that two points $\mathbf{w}_i$, $\mathbf{w}_j$ are called *adjacent on* M only if they do not belong to disconnected regions of M. This leads to the following definition:

Definition 1 *Let $M \subseteq \Re^D$ be a given manifold and $S = \{\mathbf{w}_1, ..., \mathbf{w}_N\}$ be a set of points $\mathbf{w}_i \in M$. The Voronoi polyhedra of S are denoted by V_i, $i = 1, ..., N$. Two points $\mathbf{w}_i, \mathbf{w}_j \in M \subseteq \Re^D$ are adjacent on M if their masked Voronoi polyhedra $V_i^{(M)} = V_i \cap M$, $V_j^{(M)} = V_j \cap M$ are adjacent, i.e., if $V_i^{(M)}$ and $V_j^{(M)}$ share an element $\mathbf{v} \in M$ or, equivalently, if $V_i^{(M)} \cap V_j^{(M)} \neq \emptyset$ is valid.*

Each masked Voronoi polyhedron is part of the manifold M, and the set of all masked Voronoi polyhedra forms a complete partitioning of the manifold M; i.e., $M = \bigcup_{i=1}^{N} V_i^{(M)}$ is valid.

With this definition the term "topology preserving feature map" can be formulated rigorously:

Definition 2 *Let G be a graph (network) with vertices (neural units) i, $i = 1, \ldots, N$ and edges (lateral connections) defined by the adjacency matrix $\mathbf{A}$, $A_{ij} \in \{0, 1\}$, $i, j = 1, ..., N$. Let*

$M \subseteq \Re^D$ be a given manifold of features $\mathbf{v} \in M$ and $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_N\}$ be a set of pointers (synaptic weight vectors) $\mathbf{w}_i \in M$, each of which is attached to a vertex i of the graph G. Let each feature $\mathbf{v}$ of the manifold M be mapped onto that vertex i whose pointer $\mathbf{w}_i$ is closest to $\mathbf{v}$. The graph G with its vertices i assigned to the locations $\mathbf{w}_i \in M$ forms a topology preserving map of M, if pointers $\mathbf{w}_i, \mathbf{w}_j$ which are adjacent on M belong to vertices i, j which are adjacent in G, and, vice versa, if vertices i, j which are adjacent in G are assigned to locations $\mathbf{w}_i, \mathbf{w}_j$ which are neighboring on M.

3. Induced Delaunay triangulations as perfectly topology preserving maps

Assuming the pointers $\mathbf{w}_1, ..., \mathbf{w}_N$ which are attached to the vertices i, $i = 1, ..., N$ of a graph G are distributed over the given manifold M. The graph G forms a topology preserving map of M, if vertices i, j and only vertices i, j whose corresponding masked Voronoi polyhedra $V_i^{(M)}$ and $V_j^{(M)}$ are adjacent are connected by an edge $i - j$ ($A_{ij} = 1$). The graphical structure which connects those and only those vertices i, j whose corresponding Voronoi polyhedra V_i and V_j are adjacent is the so-called *Delaunay triangulation*[1] [17]. Analog to the definition of the Delaunay triangulation $\mathcal{D}_S$ of a set of points $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_N\}$, which is based on the Voronoi polyhedra $V_1, ..., V_N$ of S, we define the *induced Delaunay triangulation* $\mathcal{D}_S^{(M)}$ based on the *masked* Voronoi polyhedra $V_1^{(M)}, ..., V_N^{(M)}$:

Definition 3 *Let $M \subseteq \Re^D$ be a given manifold and $S = \{\mathbf{w}_1, ..., \mathbf{w}_N\}$ be a set of points $\mathbf{w}_i \in M$. The induced Delaunay triangulation $\mathcal{D}_S^{(M)}$ of S, given M, is defined by the graph which connects two points $\mathbf{w}_i, \mathbf{w}_j$ if and only if their masked Voronoi polyhedra $V_i^{(M)}, V_j^{(M)}$ are adjacent, i.e., by the graph whose adjacency matrix $\mathbf{A}$, $A_{ij} \in \{0, 1\}$, $i, j = 1, \ldots, N$ has the properties*

$$A_{ij} = 1 \quad \Leftrightarrow \quad V_i^{(M)} \cap V_j^{(M)} \neq \emptyset. \tag{2}$$

We obtain the result that a graph G forms a perfectly topology preserving map of a feature manifold M, if and only if it is the induced Delaunay triangulation of the set S of pointers $\mathbf{w}_i \in M$. This is illustrated in Fig. 1. In (a), (b), (c), and (d) the given manifold M, which is disconnected and is the same in all four examples, is depicted by the two shaded areas. Only in (d), where the graph G is the induced Delaunay triangulation of the points $\mathbf{w}_i$, two vertices are connected by an edge if and only if their masked Voronoi polygons are adjacent. Only in (d) the graph G forms a perfectly topology preserving map of the given manifold M.

4. Competitive Hebbian rule

In the following we assume a set of neural units i, $i = 1, ..., N$ which develop *lateral connections* between each other, starting from being unconnected initially. A neural unit connects itself with another unit by developing a *synaptic link* to this unit. The lateral connections are described by a connection strength matrix $\mathbf{C}$ with elements $C_{ij} \in \Re_0^+$, $i, j = 1, ..., N$. The larger a matrix element C_{ij}, the stronger is the synaptic link from unit i to unit j. Only if $C_{ij} > 0$, we regard neural unit i as being connected with unit j. If $C_{ij} = 0$, neural unit i is *not* connected with unit j. Negative values for C_{ij} do not arise.

The basic principle which governs the change of interneural connection strength has first been formulated by Hebb [19]. According to Hebb's postulate a presynaptic unit i increases the strength of its synaptic link to a postsynaptic unit j if both units are concurrently active, i.e., if both activities do correlate. In its simplest mathematical formulation Hebb's rule is described by the equation

$$\Delta C_{ij} \propto y_i \cdot y_j, \tag{3}$$

[1] The Delaunay triangulation is an important structure in computational geometry, particularly for solving proximity problems.

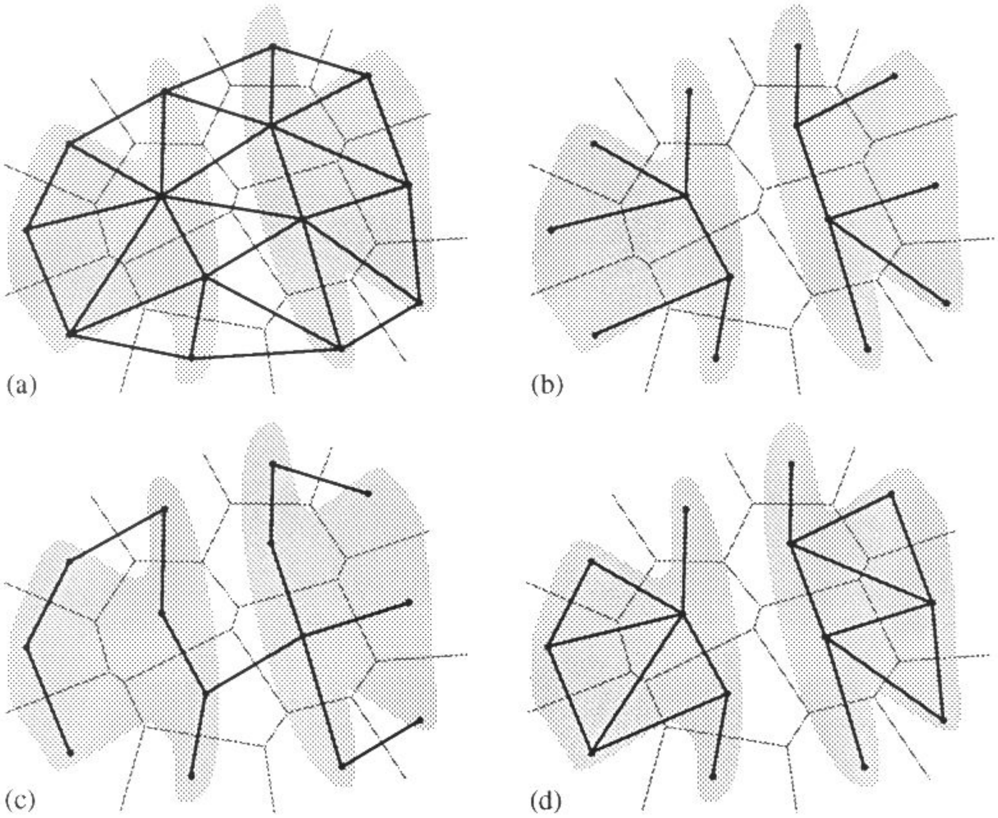

Figure 1: Illustration of our definition of topology preserving maps. In the four examples the given manifold M is disconnected and depicted by the two shaded areas. In (a) the graph G is the Delaunay triangulation of the pointers $\mathbf{w}_i$ (the location of each pointer is marked by a dot). The resulting map of M we do not consider as being topology preserving since some vertices which are connected in G belong to masked Voronoi polygons which are not adjacent on M. In (b) the opposite case is shown. The graph G does not define a topology preserving map of M since some vertices which belong to adjacent masked Voronoi polygons are not connected in G. In (c) the graph G is the minimum spanning tree of the points $\mathbf{w}_i$, as it would be suggested by the approach of Kangas et al. [15]. In (d) the graph G is the induced Delaunay triangulation. Pointers and only pointers the Voronoi polygons of which are adjacent *on* M, i.e., pointers the *masked* Voronoi polygons of which are adjacent, are connected. Only in this case the graph G forms a perfectly topology preserving map of the given manifold M.

in which the change of the strength C_{ij} of the synaptic link from unit i to unit j is linearly proportional to the presynaptic activity y_i and to the postsynaptic activity y_j. The quantities y_i, $i = 1, \dots, N$ denote the output activities of the neural units i.

We will employ the Hebb rule in a form which incorporates the novel aspect of *competition* among the synaptic links. Again to each neural unit i a weight vector $\mathbf{w}_i \in \Re^D$ is assigned. Further, we assume that each neural unit i, $i = 1, \dots, N$ receives the same afferent input patterns $\mathbf{v} \in \Re^D$. The weight vector $\mathbf{w}_i$ determines the center of the receptive field of unit i in the sense that with the reception of an input pattern $\mathbf{v}$ the output activity y_i of unit i is the larger the closer its $\mathbf{w}_i$ is to $\mathbf{v}$. In mathematical terms, we assume that $y_i = R(\|\mathbf{v} - \mathbf{w}_i\|)$ is valid, with $R(.)$ being a positive and continuously monotonically decreasing function, e.g., a Gaussian.

Applying the Hebb rule in the simple form as given in eq. (3) yields the rather trivial result that each neural unit i develops connections to all the other units $j \neq i$ with lateral connection strengths C_{ij} which are simply proportional to the overlaps of the receptive fields $R(\|\mathbf{v} - \mathbf{w}_i\|)$ and $R(\|\mathbf{v} - \mathbf{w}_j\|)$. The strength of the synaptic link between two units i and j is simply monotonically and continuously decreasing with the distance between $\mathbf{w}_i$ and $\mathbf{w}_j$. However, as in many systems governed by self-organizing processes, also the connectivity pattern which evolves on the set of neural units becomes significantly more structured if we introduce competition. In a winner-take-all network, for example, the units compete with each other based on their output activities, which finally leads to an adaptation only of the weights of the unit with the highest output activity. Without competition, all the units would behave alike and no specialization of the units, as it is characteristic for winner-take-all networks, would evolve.

Analog to the competition among the units in a winner-take-all network we introduce competition among the synaptic links. Instead of being based on the output activities of the neural units itself as in a winner-take-all network, in our model the competition among the synaptic links is determined by the *correlated* output activities Y_{ij}, the correlations of the output activities of all pairs of pre- and postsynaptic units. In the quantitative formulation given below, the correlated output activities are determined by $Y_{ij} = y_i \cdot y_j$, according to the Hebb rule (3). Keeping the analogy to winner-take-all networks, with the presentation of an input pattern $\mathbf{v}$ only the synaptic link $i - j$ whose "activity" $Y_{ij} = y_i \cdot y_j$ is highest is modified. Instead of changing the connection strengths C_{ij} according to the Hebb rule (3), in the following we will employ a winner-take-all or competitive version of (3), determined by

$$\Delta C_{ij} \propto \begin{cases} y_i \cdot y_j & \text{if } y_i \cdot y_j \geq y_k \cdot y_l \quad \forall\, k, l = 1, \dots N \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Instead of connecting each unit with all the other units, we will show that the competitive Hebb rule (4) forms a connectivity structure among the neural units i, $i = 1, \dots, N$ which corresponds to the Delaunay triangulation of the weight vectors $\mathbf{w}_1, \dots \mathbf{w}_N$. More precisely, we will show that if we present sequentially input patterns $\mathbf{v}$ with a distribution $P(\mathbf{v})$ which has support (is nonzero) everywhere on $\Re^D$, then the elements C_{ij} of the connection strength matrix $\mathbf{C}$ obey asymptotically

$$\theta(C_{ij}(t \to \infty)) = A_{ij} \qquad i, j = 1, \dots, N \tag{5}$$

with $\theta(.)$ as the Heavyside step function and A_{ij} as the elements of the adjacency matrix $\mathbf{A}$ of the Delaunay triangulation of the points $\mathbf{w}_1, \dots \mathbf{w}_N$, for which

$$A_{ij} = \begin{cases} 1 & \text{if } V_i \cap V_j \neq \emptyset \quad (V_i, V_j \text{ are adjacent}) \\ 0 & \text{if } V_i \cap V_j = \emptyset \quad (V_i, V_j \text{ are not adjacent}) \end{cases} \tag{6}$$

is valid. V_i, V_j again denote the Voronoi polyhedra of $\mathbf{w}_i$, $\mathbf{w}_j$.

For the proof we introduce the *second order Voronoi polyhedra* V_{ij}, $i, j = 1, ..., N$. The second order Voronoi polyhedron V_{ij} is given by all the $\mathbf{v} \in \Re^D$ for which $\mathbf{w}_i$ and $\mathbf{w}_j$ are the two closest points of S, i.e., V_{ij} is defined by

$$V_{ij} = \{\mathbf{v} \in \Re^D \;\big|\; \|\mathbf{v} - \mathbf{w}_i\| \leq \|\mathbf{v} - \mathbf{w}_k\| \;\wedge\; \|\mathbf{v} - \mathbf{w}_j\| \leq \|\mathbf{v} - \mathbf{w}_k\| \;\forall k \neq i, j\}. \tag{7}$$

As V_i, also V_{ij} forms a convex polyhedron. We see from (7) that the competitive Hebb rule connects two units i, j only if $V_{ij} \neq \emptyset$ is valid. Only if $\mathbf{w}_i$, $\mathbf{w}_j$ are the two points which are closest to the presented input pattern $\mathbf{v}$, $Y_{ij} = y_i \cdot y_j$ is the highest correlated output activity. We will prove that $V_{ij} \neq \emptyset$ is valid if and only if the corresponding first order Voronoi polyhedra V_i, V_j are adjacent, i.e., if and only if $V_i \cap V_j \neq \emptyset$. Then, in case $\int_{V_{ij}} P(\mathbf{v}) d\mathbf{v} \neq 0$ holds for each $V_{ij} \neq \emptyset$, the connections generated by the competitive Hebb rule form the Delaunay triangulation of the set $\mathbf{w}_1, \dots, \mathbf{w}_N$ [2].

Theorem 1 *For a set $S = \{\mathbf{w}_1, \dots, \mathbf{w}_N\}$ of points $\mathbf{w}_i \in \Re^D$ the relation*

$$V_i \cap V_j \neq \emptyset \;\Leftrightarrow\; V_{ij} \neq \emptyset \tag{8}$$

is valid. V_i denotes the first order Voronoi polyhedron of point $\mathbf{w}_i$, and V_{ij} denotes the second order Voronoi polyhedron of the points $\mathbf{w}_i$, $\mathbf{w}_j$.

[2]The following theorem and its proof has been formulated together with Philippe Dalger and Bennoit Noël [20].

Proof: If $V_i \cap V_j \neq \emptyset$ is valid, there is a $\mathbf{v} \in \Re^D$ with $\mathbf{v} \in V_i$ and $\mathbf{v} \in V_j$. Then we obtain $\|\mathbf{v} - \mathbf{w}_i\| = \|\mathbf{v} - \mathbf{w}_j\| \leq \|\mathbf{v} - \mathbf{w}_k\|$ for all $\mathbf{w}_k \in S$ and, therefore, $\mathbf{v} \in V_{ij}$, i.e., $V_{ij} \neq \emptyset$, is valid. The reverse implication follows by contradiction if we assume $V_{ij} \neq \emptyset$ and $V_i \cap V_j = \emptyset$ being valid. For each $\mathbf{v} \in V_{ij}$ the points $\mathbf{w}_i$ and $\mathbf{w}_j$ are the two nearest neighbors. Without loss of generality we can assume that for all $\mathbf{v} \in V_{ij}$ the point $\mathbf{w}_i$ is the nearest neighbor. Otherwise, against our assumption, there would be a $\mathbf{v} \in V_{ij}$ for which $\|\mathbf{v} - \mathbf{w}_i\| = \|\mathbf{v} - \mathbf{w}_j\|$ and, therefore, also $\mathbf{v} \in V_i \cap V_j$ were valid. Then, $V_{ij} \subseteq V_i$ follows. V_{ij} is given by all the $\mathbf{v} \in V_i$ which are closer to $\mathbf{w}_j$ than to all the other $\mathbf{w}_k \in S/\{\mathbf{w}_i, \mathbf{w}_j\}$. Hence, V_{ij} is bounded by hyperplanes, each of which is perpendicular to the connecting line between $\mathbf{w}_j$ and the respective $\mathbf{w}_k \in S/\{\mathbf{w}_i, \mathbf{w}_j\}$. For each hyperplane, $\mathbf{w}_j$ belongs to the half space which contains V_{ij}. Hence, $\mathbf{w}_j \in V_{ij}$ and, therefore, $\mathbf{w}_j \in V_i$ is valid. However, since also $\mathbf{w}_j \in V_j$, we obtain $\mathbf{w}_j \in V_i \cap V_j$ which is a contradiction to our assumption.

In the following we will consider pattern distributions $P(\mathbf{v})$ which have support not on the entire embedding space $\Re^D$, but only on a submanifold M. In these cases for some $V_{ij} \neq \emptyset$ the integral $\int_{V_{ij}} P(\mathbf{v}) d\mathbf{v}$ might vanish, with the result that the edge $i - j$ will not be established by the competitive Hebb rule. In these cases the competitive Hebb rule does not form the entire Delaunay triangulation, but only a subgraph of it.

5. Competitive Hebbian rule forms induced Delaunay triangulations

The competitive Hebb rule (4) constructs the full Delaunay triangulation of a set of points $\mathbf{w}_1, \ldots, \mathbf{w}_N$ only if each Voronoi polyhedron of second order V_{ij} is, at least partially, covered by the density distribution $P(\mathbf{v})$. If we define a given feature manifold M as being the manifold of $\Re^D$ on which $P(\mathbf{v})$ is non-zero, two units i, j become connected if and only if $V_{ij} \cap M \neq \emptyset$. Hence, if the manifold M forms a submanifold which does not cover each Voronoi polyhedron of second order, the Delaunay triangulation will evolve only partly. If the distribution of the points $\mathbf{w}_i$ is dense on M in a sense we will define below, the subgraph of the Delaunay triangulation which is formed by the competitive Hebb rule will be the *induced* Delaunay triangulation which was introduced in Section 3.

Definition 4 *Let $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_N\}$ be a set of points $\mathbf{w}_i$ which are distributed over a given manifold $M \subseteq \Re^D$. The distribution of the points $\mathbf{w}_i \in M$, $i = 1, ..., N$, is dense on M, if for each $\mathbf{v} \in M$ the triangle $\triangle(\mathbf{v}, \mathbf{w}_{i_0}, \mathbf{w}_{i_1})$ formed by the point $\mathbf{w}_{i_0}$ which is closest to $\mathbf{v}$, the point $\mathbf{w}_{i_1}$ which is second closest to $\mathbf{v}$, and $\mathbf{v}$ itself lies completely on M, i.e., if $\triangle(\mathbf{v}, \mathbf{w}_{i_0}, \mathbf{w}_{i_1}) \subseteq M$ is valid.*

A distribution of points $\mathbf{w}_i$ is dense on M according to the above definition, if the distribution is dense, in the common sense, compared to the topological structure of M. The distribution of the points $\mathbf{w}_i$ has to have a density which resolves the details of the shape of the submanifold M. If for each sample point $\mathbf{v} \in M$ there is a closest point $\mathbf{w}_{i_0}$ and a second closest point $\mathbf{w}_{i_1}$ such, that the triangle $\triangle(\mathbf{v}, \mathbf{w}_{i_0}, \mathbf{w}_{i_1})$ lies completely on M, the distribution of the $\mathbf{w}_i$ is dense on M. If the distribution is homogeneous, the distribution becomes dense simply by increasing the number N of points $\mathbf{w}_i$.

With Definition 4 we obtain the main theorem:

Theorem 2 *Let i, $i = 1, \ldots, N$ be a set of vertices (neural units). Let $M \subseteq \Re^D$ be a given manifold of features $\mathbf{v} \in \Re^D$ and $S = \{\mathbf{w}_1, \ldots, \mathbf{w}_N\}$ be a set of pointers (synaptic weight vectors) $\mathbf{w}_i \in M$, each of which is attached to the corresponding vertex (neural unit) i and defines the center of the receptive field $R(\|\mathbf{v} - \mathbf{w}_i\|)$ of i. If the distribution of the pointers $\mathbf{w}_i \in M$ is dense on M, then the edges (lateral connections) $i - j$ which are formed by the competitive Hebb rule define a graph (network) G which corresponds to the induced Delaunay triangulation $\mathcal{D}_S^{(M)}$ of S and, hence, forms a perfectly topology preserving map of M.*

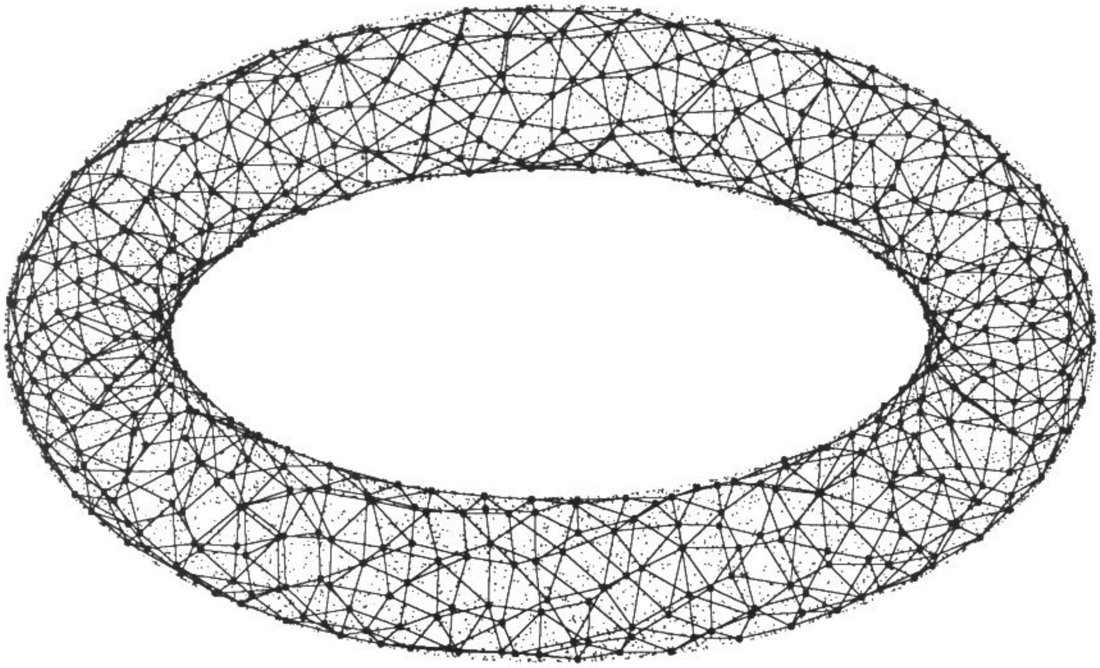

Figure 2: A topology preserving map of a torus, formed by the competitive Hebb rule. The pointers, the locations of which are marked by the large dots, were distributed over the given manifold M, i.e., the torus, by the "neural gas" algorithm [18, 21] in a preprocessing stage. Then the pointers stay fix and the edges are formed by the competitive Hebb rule. The small dots depict already presented patterns $\mathbf{v} \in M$. The few edges of the induced Delaunay triangulation which are still missing have small masked Voronoi polyhedra of second order and would emerge if further input patterns were presented.

Proof: Analog to Theorem 1 we prove the above theorem by showing that

$$V_i^{(M)} \cap V_j^{(M)} \neq \emptyset \;\Leftrightarrow\; V_{ij}^{(M)} \neq \emptyset \tag{9}$$

is valid, with $V_{ij}^{(M)} = V_{ij} \cap M$ as the masked Voronoi polyhedron of second order.

If $V_i^{(M)} \cap V_j^{(M)} \neq \emptyset$ is valid, there is a $\mathbf{v} \in M$ with $\mathbf{v} \in V_i$ and $\mathbf{v} \in V_j$. Then we obtain $\|\mathbf{v} - \mathbf{w}_i\| = \|\mathbf{v} - \mathbf{w}_j\| \leq \|\mathbf{v} - \mathbf{w}_k\|$ for all $\mathbf{w}_k \in S$ and, therefore, $\mathbf{v} \in V_{ij}^{(M)}$.

If $V_{ij}^{(M)} \neq \emptyset$ is valid, there is a $\mathbf{v}^* \in V_{ij}^{(M)}$ with $\triangle(\mathbf{v}^*, \mathbf{w}_i, \mathbf{w}_j) \subseteq M$, since the distribution of the pointers $\mathbf{w}_1, ..., \mathbf{w}_N$ is dense on M. For each $\mathbf{v} \in V_{ij}^{(M)}$ the points $\mathbf{w}_i$ and $\mathbf{w}_j$ are the two nearest neighbors. Without loss of generality we assume that for $\mathbf{v}^*$ the point $\mathbf{w}_i$ is the nearest neighbor. Since for each $\mathbf{u} \in \overline{\mathbf{v}^*\mathbf{w}_j}$ the point $\mathbf{w}_j$ is either the nearest or the second nearest neighbor of $\mathbf{u}$, and since for $\mathbf{u} = \mathbf{v}^*$ the point $\mathbf{w}_i$ is closest and for $\mathbf{u} = \mathbf{w}_j$ the point $\mathbf{w}_j$ is closest to $\mathbf{u}$, there is a $\mathbf{u}^* \in \overline{\mathbf{v}^*\mathbf{w}_j}$ for which $\|\mathbf{u}^* - \mathbf{w}_i\| = \|\mathbf{u}^* - \mathbf{w}_j\|$ is valid. Hence, we obtain $\mathbf{u}^* \in V_i$, $\mathbf{u}^* \in V_j$, and $\mathbf{u}^* \in \triangle(\mathbf{v}^*, \mathbf{w}_i, \mathbf{w}_j) \subseteq M$, which yields $M \cap (V_i \cap V_j) \neq \emptyset$ or, equivalently, $V_i^{(M)} \cap V_j^{(M)} \neq \emptyset$.

With the theorem above we have shown that the competitive Hebb rule forms perfectly topology preserving maps, supposed the distribution of the points $\mathbf{w}_i$ is dense on the given feature manifold M. In Fig. 2 we show a simulation example. The manifold M is a torus. To obtain $\mathbf{w}_i \in M$ for each i, $i = 1, ..., N$, the pointers $\mathbf{w}_i$ are distributed over M in a preprocessing stage, e.g., by a pattern driven vector quantization procedure like the "neural gas" algorithm [18, 21], which leads to a homogeneous distribution of the pointers $\mathbf{w}_i$ on M. After having distributed the pointers, the connectivity structure is formed by the competitive Hebb rule. Simply by sequentially presenting patterns $\mathbf{v} \in M$ and each time connecting those two units i, j which have the highest correlated output activity $Y_{ij} = y_i \cdot y_j$, a connectivity structure evolves which defines a perfectly topology preserving map and reflects the dimensionality and topological structure of the manifold M, i.e., of the torus.

5. Discussion

We showed how the term "topology preserving map" can be defined rigorously based on *masked Voronoi polyhedra* and *induced Delaunay triangulations*. Both the masked Voronoi polyhedra as well as the induced Delaunay triangulation of a set of points depend on the shape of the given feature manifold M. We showed that the induced Delaunay triangulation $\mathcal{D}_S^{(M)}$ as a particular subgraph of the full Delaunay triangulation $\mathcal{D}_S$ forms a perfectly topology preserving map of the manifold M. We proved rigorously and demonstrated through a computer simulation that a competitive version of the Hebb rule forms induced Delaunay triangulations and, hence, yields perfectly topology preserving maps of feature manifolds. Necessary is a distribution of the receptive field centers $\mathbf{w}_i$ of the neural units i which is dense enough to resolve the shape of the manifold M. If the manifold $M \subseteq \Re^D$ fills the embedding space $\Re^D$ completely, then the competitive Hebb rule forms the full Delaunay triangulation $\mathcal{D}_S$ as a perfectly topology preserving map of M. If M is only a submanifold of $\Re^D$, then the competitive Hebb rule forms a subgraph of $\mathcal{D}_S$, i.e., the induced Delaunay triangulation $\mathcal{D}_S^{(M)}$, as a perfectly topology preserving map of M.

Acknowledgement

I am indebted to Klaus Schulten for many fruitful discussions and for supporting this work. I thank Benoit Noël and Philippe Dalger for pointing out the connection between the competitive Hebb rule and second order Voronoi polyhedra.

References

[1] Knudsen EI, du Lac S, Esterly SD (1987), Ann. Rev. Neurosci., 10:41-65.

[2] Kohonen T (1990), Proceedings of the IEEE, 78:1464-1480.

[3] Ritter H, Martinetz T, Schulten K (1992), Neural Computation and Self-Organizing Maps. Addison-Wesley, Reading, Mass., 1992.

[4] Hubel DH, Wiesel TN (1974), J. Comp. Neurol., 158:267-294.

[5] Kaas JH, Nelson RJ, Sur M, Lin CS, Merzenich MM (1979), Science, 204:521-523.

[6] Suga N, O'Neill WE (1979), Science, 206:351-353.

[7] Kohonen T, Mäkisara K, Saramäki T (1984), Proc. 7th Int. Conf. on Pat. Recog., Montreal, pp 182-185.

[8] Brandt WD, Behme H, Strube HW (1991), Fortschritte zur Akustik– DAGA 91, Bad Honnef, DPG-GMBH, Germany, pp 1057–1060.

[9] Nasrabadi NM, Feng Y (1988), IEEE Int. Conf. on Neural Networks, San Diego, 1988, pp 1101-1108.

[10] Ritter H, Martinetz T, Schulten K (1989), Neural Networks 2(3):159-168.

[11] Willshaw DJ, von der Malsburg C (1976), Proc R Soc London B 194:431-445.

[12] Takeuchi A, Amari S (1979), Biological Cybernetics, 35:63-72.

[13] Kohonen T (1982), Biological Cybernetics, 43:59-69.

[14] Durbin R, Mitchison G (1990), Nature, 343:644-647.

[15] Kangas JA, Kohonen TK, Laaksonen JT (1990), IEEE Transactions on Neural Networks 1(1):93-99.

[16] Fritzke B (1991), Artificial Neural Networks, vol I, Kohonen T. et al. (eds.), North Holland, Amsterdam, pp. 403-408.

[17] Delaunay B (1934), Bull. Acad. Sci. USSR (VII), Classe Sci. Mat. Nat., pp 793-800.

[18] Martinetz T, Schulten K (1991), Artificial Neural Networks, vol I, Kohonen T. et al. (eds.), North Holland, Amsterdam, pp. 397-402.

[19] Hebb D (1949), Organization of Behavior. Wiley, New York.

[20] Dalger P, Noël B, Martinetz T, Schulten K (1992), Beckman Institute Technical Report TB 98-01, University of Illinois at Urbana-Champaign.

[21] Martinetz T, Berkovich S, Schulten K (1992), IEEE Transactions on Neural Networks (in press).