

Enhancing Topology Preservation during Neural Field Development via Wiring Length Minimization

Claudius Gläser, Frank Joublin, Christian Goerick

2008

Preprint:

This is an accepted article published in Proceedings of the 18th International Conference on Artificial Neural Networks - ICANN 2008, Part I. The final authenticated version is available online at: https://doi.org/[DOI not available]

Enhancing Topology Preservation during Neural Field Development Via Wiring Length Minimization

Claudius Gläser, Frank Joublin, and Christian Goerick

Honda Research Institute Europe Carl-Legien-Str. 30, D-63073 Offenbach/Main, Germany claudius.glaeser@honda-ri.de

Abstract. We recently proposed a recurrent neural network model for the development of dynamic neural fields [1]. The learning regime incorporates homeostatic processes, such that the network is able to selforganize and maintain a stable operation mode even in face of experiencedriven changes in synaptic strengths. However, the learned mappings do not necessarily have to be topology preserving. Here we extend our model by incorporating another mechanism which changes the positions of neurons in the output space. This algorithm operates with a purely local objective function of minimizing the wiring length and runs in parallel to the above mentioned learning process. We experimentally show that the incorporation of this additional mechanism leads to a significant decrease in topological defects and further enhances the quality of the learned mappings. Additionally, the proposed algorithm is not limited to our network model; rather it can be applied to any type of self-organizing maps.

1 Introduction

Self-organizing maps (SOMs) describe a group of methods in the domain of artificial neural networks. They provide learning algorithms for mapping highdimensional input data onto a (discretized) low-dimensional output space, thereby pursueing two different goals. Firstly, SOMs perform vector quantization and therewith aim at minimizing the quantization error. Secondly, they strive for achieving topology preserving mappings by incorporating topological constraints into the vector quantization process.

For technical applications the topology preserving properties of SOMs are of particular interest. They have been used in the domains of data compression [2], medical diagnosis [3], or the monitoring of industrial processes [4]. Unfortunately, SOMs tend to privilege the minimization of the quantization error over topology preservation. For this reason, various mechanisms for improving topology preservation has been suggested [5,6,7].

Dynamic neural fields are special types of SOMs, which are particularly suited for modeling cortical development [8]. However, their use in technical applications is very limited, which is mainly due to stability issues. In order to circumvent this problem, we recently proposed a recurrent neural network model for the

V. Kurkova, R. Neruda, and J. Koutnik (Eds.): ICANN 2008, Part I, LNCS 5163, pp. 593–602, 2008. © Springer-Verlag Berlin Heidelberg 2008



Fig. 1. The structure of the recurrent neural network model

homeostatic development of dynamic neural fields [1]. It is able to self-organize while maintaining a stable operation mode. However, the mappings learned by the model do not necessarily have to be topology preserving.

Here we will extend our model by incorporating another process exclusively dedicated to topology preservation. More precisely, we consider SOMs to be elastic nets with neighborhood relations between units being subject to change. We will show that a mechanism where units adapt their positions in the output space based on the concept of wiring length minimization is suitable for enhancing topology preservation. Furthermore, the method is not limited to our network model, rather it can be applied to any type of SOM.

The remainder is organized as follows. In section 2 we will shortly review our previously proposed network model. Next, we will propose a concept for enhancing topology preservation, compare it to existing approaches, and suggest a concrete implementation of it. After that, an experimental evaluation of the method will be given in section 4. The paper is finalized by a conclusion.

2 The Network Model

The recurrent neural network we presented in [1] is composed of excitatory units E and inhibitory units I, both being initially arranged on a 2-dimensional grid mimicking the neural tissue. The wiring of the network is show in Fig. 1. Afferent projections to excitatory units provide the input to the neural field. Furthermore, the units are laterally connected such that E-cells excite other E-cells as well as I-cells. In turn, E-cells receive inhibitory projections originating from I-cells.

The membrane potentials of excitatory and inhibitory units are denoted by the variables u and v, respectively. We will use i for specifying the unit located at position x_i of the cortical plane. The spatio-temporal evolution of the activity in the neural field following the presentation of a stimulus s can be described by the differential equations (1) and (2), where τ_E and τ_I are time constants, h_i^E and h^I are the resting potentials, and w_{ij}^* denotes the synaptic weight of a connection from unit j to unit i, where $* \in \{EE, EI, IE, EXT\}$ specifies the type of connection.

$$\tau_E \frac{du_i}{dt} = -u_i + \sum_j g(d_{ij}) \cdot w_{ij}^{EE} \cdot f(u_j) - \sum_j w_{ij}^{EI} \cdot f(v_j) + \sum_j w_{ij}^{EXT} \cdot s_j + h_i^E$$
(1)

$$\tau_I \frac{dv_i}{dt} = -v_i + \sum_j g(d_{ij}) \cdot w_{ij}^{IE} \cdot f(u_j) + h^I$$
⁽²⁾

Here, f is a monotonically increasing transfer function defining the relation between the membrane potential and the firing rate of a unit. We used a sigmoidal function with θ and γ denoting the threshold value and the gain factor, respectively. Additionally, we introduced a function g, which modulates the efficacy of excitatory lateral connections depending on the distance d_{ij} between the preand postsynaptic unit positions. The function g was chosen to follow the Gaussian distribution with a mean of 0.

$$f(z) = \frac{1}{1 + \exp\left(-\gamma(z - \theta)\right)} \tag{3}$$

Thus, we define that excitatory lateral connections between units within a local neighborhood are more efficient than those between far-distant units. It is, however, important to note that this is a fundamentally different approach from previously published network models. There the synaptic weight values of lateral connections are chosen as a function of the distance between the pre- and postsynaptic units, by which a topology on the neural tissue as well as within the feature space is defined. In contrast, we introduce a distance-dependent modulation of synaptic efficacy, but do not make any assumption on the synaptic weight values themselves. Following this argumentation, g could be understood as a physical constraint. The synapse location-dependent integration of excitatory synaptic input as it is carried out by passive dendrites is one possible interpretation [9].

A direct consequence of this is that large synaptic weight values could compensate for the distance-dependent modulation of connection efficacy. Thus, the mapping described by the neural field does not necessarily have to be topology preserving, that is nearby units having similar receptive fields. In the following we will propose an additional mechanism which improves the topology preservation of the learned mappings.

3 Topology Preservation

3.1 Existing Approaches

Several proposals have been made on how the topology preserving abilities of SOMs could be enhanced. Most of them rely on a fixed metric (fixed neighborhood relations) defined in the output space and try to dynamically adjust the width of the active neighborhood based on global heuristics or more sophisticated local measurements like input novelty [6], topology defects [5], or the degree of local folding [10]. Other approaches aim at achieving a reasonable balance between the two objective functions of minimizing the quantization error



Fig. 2. An illustrative comparison between convential SOM learning algorithms (a) and the proposed system for enhancing topology preservation of SOMs (b)

and topological defects by differently weighting their influence to the learning algorithm [7]. Only a few approaches try to change the metric defined in the output space. They mostly concentrate on building tree-like neighborhood relations by hierarchically clustering codebook vectors [11].

3.2 Topology Preservation Via Wiring Length Minimization

According to Fig. 2(a) SOMs can be considered to be a group of methods for vector quantization, insofar as the codebook vectors of their units are adapted in order to minimize the quantization error. Thereby, topological constraints, which are defined by the units' neighborhood relations, are incorporated. By doing so, SOMs aim at finding a reasonable balance between the two objectives of minimizing the quantization error and minimizing topological defects of the mapping. Nevertheless, when mapping high-dimensional input data to a low-dimensional output space, the requirements described by the two objective functions most often cannot be simultaneously satisfied. In these cases, SOMs privilege vector quantization at the cost of an increase in topological defects.

We propose to release (or at least to relax) the topological constraints from the process of vector quantization by incorporating another process which is exclusively dedicated to topology preservation. This means that the objective function of minimizing topological defects is no longer implicitly defined by incorporating topological constraints into the vector quantization method; rather it is explicitly defined by a separate process running in parallel to vector quantization. Therefore, we propose to add another process to the usual SOM learning algorithms for vector quantization which changes the positions of the units in the output space based on the result of vector quantization, such that topology preservation is improved (see Fig. 2(b)). This means that, in the output space, a fixed metric is used (e.g. Euclidean distance), but the neighborhood relations between units are changed by adjusting their positions. Therewith, SOMs can be seen as elastic networks of laterally connected units. Elastic means that individual units are placed in the output space, which is a continuous (physical) substrate with a fixed metric, and that individual units are able to dynamically adjust their positions on this substrate. More precisely, units are interconnected in a way, that the strength w_{ij} of the connection between two units *i* and *j* denotes the similarity between the codebook vectors w_i and w_j of the two units.

$$w_{ij} \propto \frac{1}{\|w_i, w_j\|^I} \tag{4}$$

There $\|\cdot\|^I$ is the distance metric of the input space I.

We propose that the process of adjusting the position of units in the output space is based on an individual unit's objective function of minimizing the weighted wiring length of its connections with other units of the SOM. Let x_i denote the position of a unit *i* in the output space and $\|\cdot\|^A$ be the distance metric of the output space. Then the objective function of a unit *i* for changing its position in the output space can be formulated according to (5), which is a purely local objective function. Alternatively, the problem can be formulated using a global objective function according to (6), which states to minimize the total weighted wiring length of the elastic network.

$$\sum_{j} w_{ij} \cdot \parallel x_i, x_j \parallel^A \longrightarrow \min$$
 (5)

$$\sum_{i} \sum_{j} w_{ij} \cdot \parallel x_i, x_j \parallel^A \quad \longrightarrow \quad \min \tag{6}$$

3.3 Implementation of Wiring Length Minimization

If we apply the above mentioned framework for topology preservation via wiring length minimization to our network model described in section 2, we first have to chose appropriate values for the connection strengths of the elastic network. One possibility would be to chose them according to (4). However, in the case of developing neural fields with learned lateral connections as those featured by our network model, one can directly use the synaptic weight values of the within field connections. This is possible, because the synaptic weights of lateral connections are a measure for the similarity between the receptive fields of different units. If we do so, we obtain the following objective functions of individual units, where (7) and (8) hold for an excitatory unit *i* and an inhibitory unit *i*, respectively. Furthermore, d_{ij}^* with $* \in \{EE, EI, IE\}$ denotes the distance between units *i* and *j* according to the distance metric of the output space $\|\cdot\|^A$.

$$\sum_{j} w_{ij}^{EE} \cdot d_{ij}^{EE} + \sum_{j} w_{ji}^{EE} \cdot d_{ji}^{EE} + \sum_{j} w_{ji}^{IE} \cdot d_{ji}^{IE} \longrightarrow \min$$
(7)

$$\sum_{j} w_{ij}^{IE} \cdot d_{ij}^{IE} + \sum_{j} w_{ji}^{EI} \cdot d_{ji}^{EI} \longrightarrow \min$$
(8)



Fig. 3. The attractive and repulsive forces exerted on model units depend on the strengths of the connections as well as the distances between units

Now, multiple optimization techniques could be applied in order to minimize the wiring length. In example unit positions could be adapted via gradient descent or evolutionary strategies. Here, we follow another approach in which we interpret connections between units as springs with spring constants chosen according to the strengths w_{ij} of the connections. Following this argumentation, connections of the elastic net exert attractive forces on units in a way that the force $F_{i,ij}^+$ exerted on a unit *i* by a connection between individual units *i* and *j* increases with an increase in the connection strength w_{ij} or an increase in the distance d_{ij} between the two units and vice versa.

$$F_{i,ij}^+ \propto w_{ij} \cdot d_{ij} = w_{ij} \cdot || x_i, x_j ||^A$$
(9)

A trivial solution of (7) and (8) would be to set the distances between individual units to 0. The approach outlined above would converge to this solution as well. This would mean that all units coincide at the same position in the output space. For this reason, we additionally consider repulsive forces between units of the elastic net, such that the repulsive force $F_{i,ij}^-$ exerted on a unit *i* increases when the distance d_{ij} between the units *i* and *j* decreases and vice versa. The resulting physical interaction between units is illustrated in Fig. 3.

$$F_{i,ij}^{-} \propto \frac{1}{d_{ij}} = \frac{1}{\|x_i, x_j\|^A}$$
 (10)

Additionally, it is important to note that in our model repulsive forces only exist between pairs of excitatory units as well as between pairs of inhibitory units, whereas there is no repulsive force between an excitatory and an inhibitory unit. This is reasonable when considering the output space to be the cortical tissue with excitatory and inhibitory units being located in different layers of it. Therewith the positions of excitatory and inhibitory units can be updated according to the following equations:

$$\Delta x_i^E \propto \sum_j F_{i,ij}^{+,EE} + \sum_j F_{i,ij}^{+,IE} + \sum_j F_{i,ij}^{+,EI} + \sum_j F_{i,ij}^{-,EE} \qquad (11)$$

$$\Delta x_{i}^{I} \propto \sum_{j} F_{i,ij}^{+,IE} + \sum_{j} F_{i,ij}^{+,EI} + \sum_{j} F_{i,ij}^{-,II}$$
(12)

4 Experimental Results

In order to evaluate the presented method for improving the topology preservation of neural fields, we carried out an experiment in the domain of reference frame transformation. More precisely, we investigated the use of our recurrent neural network model for 1-dimensional eye-hand coordination. In order to robustly perform eye-hand coordination, an animal has to be able to transform between the different reference frames, an ability usually attributed to an intermodal body-calibration obtained in the early stages of development. The key aspect there is that simultaneously present stimuli become linked together and can later be used for the transformation from one modality into another. Here we want to use our network for modeling the calibration process during early self-exploration. Therefore, we have chosen three stimuli s^1, s^2, s^3 with $s^1, s^2 \in [-1, 1]$ and $s^3 = s^1 - s^2$, where s^1 and s^2 mimic the gaze and hand position in a body-centered reference frame, respectively, as well as s^3 representing the hand position in eye-centered coordinates.

Each of the three stimuli s^1, s^2, s^3 is represented by a population code composed of 21 neuron responses, resulting in a total of 63 inputs to the neural field. Target gaze- and hand-positions (s^1, s^2) were chosen randomly. The recurrent neural network is composed of 100 excitatory units and 100 inhibitory units, both initially being arranged on a 10x10 grid. The synaptic weight values of afferent projections to the field w_{ij}^{EXT} were initialized with small random values. Weights of lateral connections were initialized uniformly. In order to evaluate the effect of topology preservation via wiring length minimization (WLM), we compared the results obtained by two simulation runs; one run using WLM and one run not using WLM.

Firstly, we investigated if our interpretation of a neural field to be an elastic net with lateral connections exerting attractive forces on units is a suitable choice for minimizing wiring length. Therefore, the total weighted wiring length (see (6)) was calculated and normalized by the mean distance between units. Fig. 4(a) shows the temporal evolution of this measure for two developing neural fields, one using the process of WLM and one not using WLM. As can be seen the process of WLM eliminates the initial increase in total wiring length, which can be attributed to an initial rough adjustment of synaptic weight values. Later in training, the competitive nature of the learning regime incorporated by our network model induces a "die off" of many synapses, which results in a decrease in total weighted wiring length for both simulation runs. However, the neural field using WLM converges to a smaller value, which shows that our implementation is a suitable choice for WLM.

Secondly, the way how neurons change their positions in the output space will be illustrated. For the sake of simplicity we will concentrate on excitatory units. Therefore, Fig. 4(b) shows the starting and end positions of individual units. The neuron traces of four exemplarily chosen units are additionally shown. They illustrate that some units change their positions considerably. Furthermore, the traces indicate that position changes are large during the initial training phase, whereas units maintain their positions at the end of training.



Fig. 4. (a) The total wiring length of the elastic net as a function of training time. (b) The positions of excitatory units before (open circles) and after training (filled circles) using WLM. Additionally, the traces of 4 exemplarily chosen units are shown (dotted lines).

The results presented so far focused on how WLM is achieved, however, they do not demonstrate the suitability of WLM for improving topology preservation. In order to do so we investigated the topology preserving properties of the mapping described by the neural field with a special emphasize on the effect of WLM. Therefore, we trained two neural fields: one using the learning regime incorporating WLM and one not using WLM. After training neighborhood relations were defined using Delaunay triangulation of the units' end positions. Finally, we plotted the position of the units' receptive fields (codebook vectors) in the input space (see Fig. 5 (a) and (b)), where the connections between units are plotted according to the neighborhood relations defined in output space. For a topology preserving mapping, this would result in a network, where adjacent



Fig. 5. The position of the units' receptive fields (codebook vectors) in input space as well as their neighborhood relations defined in output space for a neural field trained not using WLM (a) and one using WLM (b)



Fig. 6. The topographic function as a measure for the degree of topological defects is plotted for neural fields using WLM or not using WLM in the course of training

units in input space feature a connection according to the neighborhood relations defined in output space. As can be seen, this is the case for the neural field which was trained using WLM, whereas it is not the case when not using WLM.

In order to give a quantitative measure for the topology preserving properties of the learned mappings, we calculated the topographic function [12], which is a widely used method for characterizing the degree of topology preservation. Fig. 6 shows the corresponding plots, where the rank k determines the effective neighborhood range on which topology preservation was analyzed. More precisely, the topographic function for small |k| describes the degree of local topological defects, whereas the topographic function for large |k| describes the degree of global topological defects. Furthermore, a topographic function of 0 would be obtained when analyzing a mapping which is perfectly topology preserving. The plots of Fig. 6 show that the incorporation of WLM into the learning regime for the development of dynamic neural fields results in a significant decrease of topological defects on both a local scale and particularly on a global scale.

5 Conclusion

Our previous work [1] focused on the incorporation of recent advances in the understanding of how homeostatic processes regulate neuronal activity. There we could show that our network model is able to self-organize while maintaining a stable operation mode even in face of experience-driven changes. In this paper we extended our existing network model by another process, which enhances the topology preserving properties of the learned mappings. Therefore, we consider the neural field to be an elastic net with unit positions (and therewith neighborhood relations) being subject to change, where the objective function for changing the unit positions is the minimization of wiring length. With other words, units try to adapt their positions in the output space such that the distance to their neighbors in input space becomes minimized. For a concrete implementation we suggested to interpret connections between units as springs exerting forces on units. We experimentally showed that the incorporation of this additional mechanism for wiring length minimization significantly decreases

the topological defects of the learned mapping. Furthermore, the self-regulatory properties of our network are not affected by this additional process.

The combination of these two algorithms further improves the quality of our network model and should ease its application in various domains. On the one hand, the self-regulatory nature of the model avoids stability problems, which one usually faces when using dynamic neural fields. On the other hand, the now incorporated enhancement of topology preservation broadens the domains to which the model can be potentially applied.

Additionally, the proposed mechanism for enhancing topology preservation is not limited to our network model. It is even not limited to dynamic neural fields. Rather it can be applied to various methods in the field of self-organizing maps, i.e. the widely used Kohonen-maps [13].

References

- 1. Gläser, C., Joublin, F., Goerick, C.: Homeostatic development of dynamic neural fields. In: Proc. International Conference on Development and Learning (2008)
- Kangas, J.: Utilizing the similarity preserving properties of self-organizing maps in vector quantization of images. In: Proc. International Conference on Neural Networks, vol. 4, pp. 2081–2084 (1995)
- Vörös, T., Keresztényi, Z., Fazekas, C., Laczkó, J.: Computer aided interactive remote diagnosis using self-organizing maps. In: Proc. 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 5, pp. 3190–3193 (2004)
- Alhoniemi, E., Hollmén, J., Simula, O., Vesanto, J.: Process monitoring and modeling using the self-organizing map. Integrated Computer-Aided Engineering 6(1), 3–14 (1999)
- Herrmann, M.: Self-organizing feature maps with self-organizing neighborhood widths. In: Proc. International Conference on Neural Networks, pp. 2998–3003 (1995)
- Phaf, R., Den Dulk, P., Tijsseling, A., Lebert, E.: Novelty-dependent learning and topological mapping. Connection Science 13(4), 293–321 (2001)
- Kirk, J.S., Zurada, J.M.: Algorithms for improved topology preservation in selforganizingmaps. In: Proc. International Conference on Systems, Man, and Cybernetics, vol. 3, pp. 396–400 (1999)
- Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. Biological Cybernetics 27, 77–87 (1977)
- 9. Williams, S., Stuart, G.: Role of dendritic synapse location in the control of action potential output. Trends in Neuroscience 26(3), 147–154 (2003)
- Kiviluoto, K.: Topology preservation in self-organizing maps. In: Proc. International Conference on Neural Networks, vol. 1, pp. 294–299 (1996)
- Kirk, J.S., Zurada, J.M.: A two-stage algorithm for improved topography preservation in self-organizing maps. In: Proc. International Conference on Systems, Man, and Cybernetics, vol. 4, pp. 2527–2532 (2000)
- Villmann, T., Der, R., Herrmann, M., Martinetz, T.M.: Topology preservation in self-organizing feature maps: Exact definition and measurement. IEEE Transactions on Neural Networks 8(2), 256–266 (1997)
- Kohonen, T.: Self-organized formation of topologically correct feature maps. Biological Cybernetics 43, 59–69 (1982)