Unsupervised Learning in Networks of Spiking Neurons Using Temporal Coding

Berthold Ruf and Michael Schmitt*

Institute for Theoretical Computer Science, Technische Universität Graz
Klosterwiesgasse 32/2, A-8010 Graz, Austria
E-mail: {bruf, mschmitt}@igi.tu-graz.ac.at

Abstract. We propose a mechanism for unsupervised learning in networks of spiking neurons which is based on the timing of single firing events. Our results show that a topology preserving behaviour quite similar to that of Kohonen's self-organizing map can be achieved using temporal coding. In contrast to previous approaches, which use rate coding, the winner among competing neurons can be determined fast and locally. Hence our model is a further step towards a more realistic description of unsupervised learning in biological neural systems.

1 Introduction

In the area of modelling information processing in biological neural systems, there is an ongoing debate about which essentials have to be taken into account (see e.g. [3,13,11,9]). Discrete models, such as threshold gates or McCulloch-Pitts neurons, are undoubtedly very simplistic descriptions of biological neurons. Models with real-valued output, such as the sigmoidal gate, where analogue values are interpreted as firing rates of biological neurons, are more suitable for the modelling of neural processes in terms of analogue computations. However, both types of models do not capture a phenomenon which is widely believed to be the basis of fast analogue computations in biological neural networks: the timing of single action potentials. Models of spiking neurons that use temporal coding are a first attempt to explore the computational significance of this phenomenon [1,7,11]. Furthermore, spiking neuron networks (SNNs), where the computations are based on this coding scheme, have recently been shown to be computationally more powerful than networks consisting of threshold or sigmoidal gates with respect to time and network complexity [9].

In this paper we investigate unsupervised learning processes in SNNs. On the basis of a construction introduced in [8] we show how competitive learning can be performed by SNNs using temporal coding. We extend this idea to a learning mechanism that is closely related to one of the most successful paradigms of unsupervised learning: the self-organizing map (SOM) by Kohonen [6].

Topology preserving maps have been found in many regions of the brain, e.g. in the visual, auditory, or somatosensory cortex [1]. The SOM provides a

* To whom correspondence should be addressed.

possible explanation how such maps can develop. Previous versions of the SOM assume that the output of a neuron is characterized by its firing rate and not by the timing of single firing events. Using lateral connections the procedure for detecting the so-called winner neuron is usually implemented as a recurrent network. This has the consequence that the winner neuron is detected not before the network has settled down into an equilibrium state. However, since the computation relies on the convergence of the network, this approach disregards the benefits that temporal coding offers to fast information processing in SNNs.

In addition to these conventional implementations there has also been some research on biologically more realistic models of self-organizing map algorithms, e.g. by Kohonen [5], Sirosh and Miikkulainen [14], Choe and Miikkulainen[2]. Also in these approaches the output of a neuron is assumed to correspond to its firing rate, and learning takes place in terms of this rate after the network has reached a stable state of firing.

In the following we propose a mechanism for unsupervised learning in SNNs where computing and learning are based on the timing of single firing events. In contrast to the standard formulation of the SOM, our construction has the additional advantage that the winner among the competing neurons can be determined fast and locally by using lateral excitation and inhibition. These lateral connections also constitute the neighbourhood relationship among the neurons. We assume that initially neurons which are topologically close together have strong excitatory lateral connections whereas remote neurons have strong inhibitory connections. During the learning process the lateral weights are decreased, thus reducing the size of the neighbourhood.

In a series of computer simulations we have investigated the capability of the model to form topology preserving mappings. For the evaluation of these mappings, instead of relying on visual inspection, we used a measure for quantifying the neighbourhood preservation which was recently studied [4]. Our results show that the model exhibits the same characteristic behaviour as the SOM. The typical emergence of topology preserving behaviour could be observed for a wide range of parameters.

We also studied the effect of weight normalization. It is used in several implementations of the SOM (see e.g. [2,6,14]) but its biological relevance is controversial. We performed simulations with and without weight normalization after each learning cycle. Comparing the results we discovered that after a certain number of learning cycles approximately the same degree of topology preservation could be achieved regardless whether the weights were normalized or not. The model of unsupervised learning that we propose in this paper is therefore a candidate for a more realistic description of fast analogue computation in biological neural systems. Moreover, it also provides a link to possible industrial applications via silicon implementations in pulse coded VLSI [10].

The paper is organized as follows: In Section 2 we introduce the formal model of a spiking neuron. In Section 3 we propose the mechanism for unsupervised learning in networks of these model neurons. The simulation results are presented in Section 4.

2 The Spiking Neuron Model

Recently Maass has shown in [8] how leaky integrate-and-fire neurons can compute weighted sums in temporal coding, where the firing time of a neuron encodes a value in the sense that an early firing of the neuron represents a large value. More precisely, one considers a neuron v which receives excitatory input from m neurons $u_1, \ldots, u_m$, where the corresponding weights are denoted by $w_1, \ldots, w_m$. Each u_i fires exactly once within a sufficiently small time interval at a time t_i with $t_i = T_1 - s_i$, where s_i is the ith input to v and T_1 some constant. Usually T_1 is given by the time when a reference spike is fired by some additional input neuron to v. Under certain weak assumptions (basically the initial rising segments of the excitatory postsynaptic potentials have to be linear) one can guarantee that v fires at a time determined by $T_2 - \sum w_i s_i$ with T_2 being some constant.

This computation can also be performed on the basis of "competitive temporal coding", such that no explicit reference times T_1 and T_2 are necessary (see [8] for details).

3 Unsupervised Learning

On the basis of this construction it is now possible to implement various types of unsupervised learning in the context of SNNs as follows: Given a set S of m-dimensional input vectors $\mathbf{s}^l = (s_1^l, \ldots, s_m^l)$ and an SNN with m input neurons and n output neurons, where each output neuron v_j receives synaptic "feedforward" input from each input neuron u_i with weight w_{ij} and "lateral" synaptic input from each output neuron $v_k, k \neq j$, with weight $\tilde{w}_{kj}$. At every cycle of the learning procedure one $\mathbf{s}^l \in S$ is randomly chosen and the input neurons are made fire such that they temporally encode $\mathbf{s}^l$. Each v_j then starts to compute $\sum_i w_{ij} s_i^l$ as described in Section 2. If we assume that the input vector and the weight vector for each neuron are normalized, then this weighted sum represents the similarity between the two vectors with respect to the euclidean distance. Hence the earlier v_j fires, the more similar is its weight vector to the input vector and the output neurons form a competitive layer, where the winner is the neuron which fires first. Note that the firing time of the winner is not influenced by the firing of the other output neurons.

If the lateral connections are strongly inhibitory, such that the firing of the winner neuron, say v_k, inhibits all other neurons in the competitive layer from firing, one can implement in a straightforward way competitive learning: One simply has to apply the standard competitive learning rule (also known as instar learning rule) to the winner neuron v_k. This rule is given for v_k by:

$$\Delta w_{ik} = \eta(s_i^l - w_{ik}), \qquad i \in \{1, \ldots, m\}$$

where $\mathbf{s}^l$ is the current pattern and η the learning rate.

For realizing the SOM in our model, one has to find a way to implement the neighbourhood function on the basis of locally available information. We use

the lateral connections among the output neurons to reflect the structure of the neighbourhood by assuming that initially neurons which are topologically close together have strong excitatory lateral connections whereas remote neurons have strong inhibitory connections. This means that the firing of the winner neuron, say v_k, at time t_k drives the firing times of neurons in the neighbourhood of v_k towards t_k, thus increasing the values they encode. The firing of remote neurons is postponed by the lateral inhibition. We suggest the following learning rule:

$$\Delta w_{ij} = \eta \frac{T_{out} - t_j}{T_{out}} (s_i^l - w_{ij}) \tag{1}$$

where t_j is the firing time of the jth output neuron. The learning rule applies only to neurons that have fired before a certain time T_{out} (which has to be chosen sufficiently large). The factor $(T_{out} - t_j)/T_{out}$ realizes the neighbourhood function, which is largest for the winner neuron and decreases for neurons which fire at later times.

During the learning process the lateral weights $\tilde{w}_{kj}$ are decreased, thus reducing the size of the neighbourhood.[1] As in the standard formulation of the SOM, η is slowly decreased during learning.

4 Simulations

We tested our approach with one of the standard examples for the SOM, where two-dimensional input patterns are chosen randomly from a square and the competitive units are expected to organize themselves in a topology preserving grid. For the tests we used an array of 5×5 competitive units. We normalized the inputs by adding a third input component, which was chosen such that all input vectors have the same norm. The feedforward weights were initialized from random values around the midpoint of the patterns. The lateral weights for the immediate neighbours were chosen slightly positive, for the second neighbours zero and for all other neurons negative. In order to examine the effect of normalization we performed two series of experiments. In the first the feedforward weights were normalized after each application of (1), in the second they remained unnormalized.

In order to quantify the neighbourhood preservation we used the method of "metric multidimensional scaling" (see e.g. [4]) which is based on the measure

$$M_{\mathrm{MDS}} = \sum_{i=1}^{N} \sum_{j<i} (F(i,j) - G(M(i), M(j)))^2. \tag{2}$$

where $N = |S|$. The function M represents the mapping of the network, i.e. $M(i)$ is the index of the winner neuron in the competitive layer for input $\mathbf{s}^i$. The

[1] This requires that the lateral weights change their sign during learning, which is not very realistic in the context of biological networks. However, one can consider instead two connections, one excitatory, which decreases and one inhibitory which increases during learning.

function F specifies for any given pair of patterns how dissimilar they are, and G is the measure of dissimilarity between pairs of weight vectors. Obviously, a value of $M_{\mathrm{MDS}} = 0$ indicates a perfectly topology preserving mapping M. In order to take the random initialization of the weights into account we computed the relative neighbourhood distortion which is the actual value of M_{MDS} divided by the maximum initial value of M_{MDS}. This yields an initial value from the interval $[0, 1]$ and makes the results for different initializations comparable. Figure 1 shows that – regardless whether the weights are normalized after each cycle or not – the relative M_{MDS} stabilizes for our example indeed at approximately the same small value.

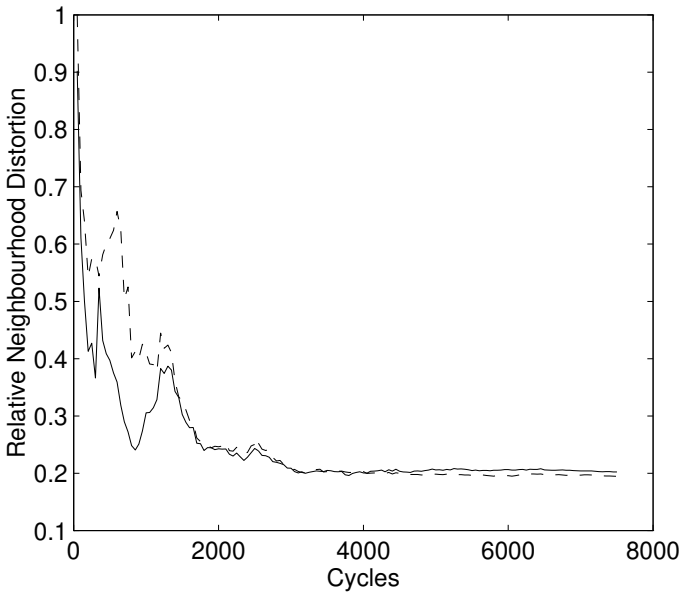

Fig. 1. Development of the relative neighbourhood distortion for the example described in Section 4, where the solid (dashed) line shows the case with (without) weight normalization after each cycle.

5 Conclusions

To the best of our knowledge, this is the first implementation of Kohonen's learning algorithm for SNNs using temporal coding. This work is a further step towards showing that biological neurons can indeed achieve a topology preserving behaviour using a similar learning procedure like the one suggested by Kohonen.

As mentioned above we have assumed as in [8] that the neurons are of the leaky integrate-and-fire type, where the initial segment of the postsynaptic potentials rises linearly. Recent simulations [12] have shown that even when using a more detailed neural model which includes nonlinear effects and more realistic shapes (e.g. α-functions) for the postsynaptic potentials, spiking neurons can still compute weighted sums in the way described in Section 2. This indicates that these simplifying assumptions can be dropped.

References

1. Arbib, M. A.: The Handbook of Brain Theory and Neural Networks. MIT Press, Cambridge (1995).
2. Choe, Y., Miikkulainen, R.: Self-organization and segmentation with laterally connected spiking neurons. Technical Report AI TR 96-251, Department of Computer Science, University of Texas at Austin, September 1996.
3. Gerstner, W., van Hemmen, L. H.: How to describe neuronal activity: spikes, rates, or assemblies? In Advances in Neural Information Processing Systems 6, Morgan Kaufmann, San Mateo (1994) 463–470.
4. Goodhill, G. J., Sejnowski, T. J.: Quantifying neighbourhood preservation in topographic mappings. Proceedings of the 3rd Joint Symposium on Neural Computation, San Diego, CA (1996) 61–82.
5. Kohonen, T.: Physiological interpretation of the self-organizing map algorithm. Neural Networks **6** (1993) 895–905.
6. Kohonen, T.: Self-organizing maps. Springer Verlag, Berlin (1995).
7. Maass, W.: Lower bounds for the computational power of networks of spiking neurons. Neural Computation **8** (1996) 1–40.
8. Maass, W.: Fast sigmoidal networks via spiking neurons. Neural Computation **9** (1997) 279–304.
9. Maass, W.: Networks of spiking neurons: the third generation of neural network models. Neural Networks, to appear; extended abstract in Proceedings of the Seventh Australian Conference on Neural Networks, Canberra (1996) 1–10; available from ftp-host: archive.cis.ohio-state.edu, ftp-filename: /pub/neuroprose/maass.third-generation.ps.Z.
10. Murray, A., Tarassenko, L.: Analogue Neural VLSI: A Pulse Stream Approach. Chapman & Hall, London (1994).
11. Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W.: SPIKES: Exploring the Neural Code. MIT Press, Cambridge (1996).
12. Ruf, B.: Computing functions with networks of spiking neurons. To appear in: Proceedings of the International Work-Conference on Artificial and Biological Neural Networks (1997).
13. Sejnowski, T.: Time for a new neural code? Nature **376** (1995) 21–22.
14. Sirosh, J., Miikkulainen, R.: Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex. Neural Computation **9** (1997) 577–594.