

# Selective attention improves self-organization of cortical maps with multiple inputs

Thomas Trappenberg, Aya Saito, and Pitoyo Hartono

**Abstract**—Models of self-organizing cortical maps have focused on demonstrations with single objects in the environment. Recently, the validity of a traditional biological model has been questioned for the case of multiple simultaneous input sources. Here we show that the standard model is able to self-organize with multiple inputs. However, we also show that the ability to self-organization can be enhanced considerably by including top-down attention as well as some noise. The model is also used to simulate the development of tuning curves.

## I. INTRODUCTION

Experience-dependent formations and functional reorganizations of topographic maps have long been an important demonstration of system-level brain plasticity. A better understanding of the underlying mechanisms is relevant for developing advanced learning techniques and rehabilitation strategies. Topographic maps have been found in many sensory areas such as orientation maps in the visual cortex [1], tonotopic maps in the primary auditory cortex [2], and somatosensory maps in the somatosensory cortex [3], and sensory-driven changes in such maps have long been studied [4]. The early development of cortical map organization is often attributed to a critical period [5], but recent findings have shown that reorganizations are possible in mature brains [6]. Some of the first models that capture such plasticity-dependent organization of cortical maps have been proposed by Willshaw and von der Malsburg [7], [8], and Amari and Takeuchi [9], [10]. All of these models rely on a competitive dynamics in the map layer realized through lateral interactions. The lateral weights have been fixed in the early studies while modifying the afferent connections to the cortical maps, and the self-organization was primarily studied through the visualization of the organization of the centers of the resulting tuning curves. Such models have also been applied to higher dimensional feature maps and to map high dimensional feature spaces into lower dimensional maps [11], [12], we discuss here the case where the dimensionality of the feature space is the same as the dimensionality of the feature map.

The above models have been enhanced by Sirosh and Miiikkulainen [13], [14] in several ways (see also [15]). In particular, these authors considered learning of the lateral weights, and they included some other details such as weight

normalization which seem to make the simulations more robust. While these models are formulated with large generality, previous studies have not focused on the development of the form of the tuning curves from these first principles other than noting that the self-organization ability is strongly influenced by the size of initial anatomical mini-patches [14]. In this paper we study specifically the development of tuning curves but keep the lateral weights fixed as in the original models.

A further important model of self-organizing maps (SOM<sup>1</sup>) was proposed by Kohonen [16]. Kohonen’s model contains important abstractions and simplifications of the previous models, which makes his model applicable to large scale simulations and to many technical applications [17], [18]. While the Kohonen SOM has rightly dominated technical applications of self-organizing maps, certain aspects of biological systems can only be studied in more detailed models. For example, in the biological interpretation of Kohonen maps, tuning curves are assumed to be fixed (usually Gaussian). Although the basic Kohonen model has been enhanced by considering different kernel functions [19], [20], [21], we consider here the development of tuning curves here from the first principles based on the physiological models of self-organizing maps.

In contrast to the original models of self-organizing maps, Kohonen’s model is applicable only in situations where one stimulus is dominating the input to a cortical map. In order to validate the self-organizing principles in biological systems, natural scenes that consist of multiple objects should be considered. Recently, Kohonen [22] argued that the physiological SOM model [8], [10] does not work with multiple inputs, and he proposed a modified model which he showed can organize with multiple inputs. Here we show that the basic physiological SOM model can self-organize with multiple inputs, but that self-organization is considerably enhanced by taking top-down mechanisms, such as attention, into account. Furthermore, we show that such top-down mechanisms are essential when supervising the learning of specific actions based on self-organized representation in the cortex.

## II. THE MODEL AND EXPERIMENTAL SETUP

We consider a dynamical neural field (DNF) model of a cortical sheet [9], where the activation  $\mathbf{r}(\mathbf{x}, t)$  of a localized neural assembly is determined by an internal state variable  $\mathbf{u}(\mathbf{x}, t)$ , which is governed by a leaky integrator dynamics

Thomas Trappenberg is with the Department of Computer Sciences and the Dalhousie Neuroscience Institute, Dalhousie University, 6050 University Avenue, Halifax NS B3H 1W5, Canada (phone: +1 902 3087; email: tt@cs.dal.ca).

Aya Saito and Pitoyo Hartono are with the School of Media Architecture, Future University-Hakodate Kamedanakanochi 116-2, Hakodate, Hokkaido 041-8655 Japan (phone +81-138-34-6462; e-mail: hartono@fun.ac.jp)

<sup>1</sup>While the acronym SOM was coined by Kohonen, we will refer to any model of a self-organizing map as SOM.

with external input  $I^{\text{ext}}(\mathbf{x}, t)$

$$\tau \frac{\partial \mathbf{u}(\mathbf{x}, t)}{\partial t} = -\mathbf{u}(\mathbf{x}, t) + \int_{\mathbf{y}} \mathbf{w}(\mathbf{x}, \mathbf{y}) \mathbf{r}(\mathbf{y}, t) d\mathbf{y} + I^{\text{ext}}(\mathbf{x}, t) \quad (2)$$

$$\mathbf{r}(\mathbf{x}, t) = \frac{1}{1 + \exp(-\beta \mathbf{u}(\mathbf{x}, t))} \quad (2)$$

$$I^{\text{ext}}(\mathbf{x}, t) = \int_{\mathbf{y}} \mathbf{w}^{\text{ext}}(\mathbf{x}, \mathbf{y}) \mathbf{r}^{\text{ext}}(\mathbf{y}, t) d\mathbf{y}. \quad (3)$$

The weights  $w$  describe the connection strength between cortical assemblies within the cortical sheet. In this paper we use a shifted Gaussian

$$\mathbf{w}(|x - y|) = A_w \left( e^{-(x-y)^2/4\sigma^2} - C \right), \quad (4)$$

where  $A_w$ ,  $\sigma$ , and  $C$  are constants, and  $x$  and  $y$  are the locations of the interacting neural populations. To minimize boundary effects, we implement the model on a torus with extension  $2\pi$ . In this case, the distances  $|x - y|$  should be replaced by  $\text{MIN}|x - y|, 2\pi|x - y|$  in the above formulas. The precise form of the weight kernel is not important for the following arguments as long as it describes short distance excitation and long distance inhibition. The above weight kernel can be learned through Hebbian learning [13], [14], [23], but most studies of self-organizing maps consider this interaction kernel as fixed when learning the afferent (input) weights,  $\mathbf{w}^{\text{ext}}$ , to this neural sheet (see [24] for a general introduction to DNF models). Amari [9] has shown that neural field models have solutions with localized activity packets, and such networks can hence function as a form of winner-takes-all (WTA) network. The winner is thereby determined not only by the node with highest activation, but depends also on the support it gets from neighboring nodes.

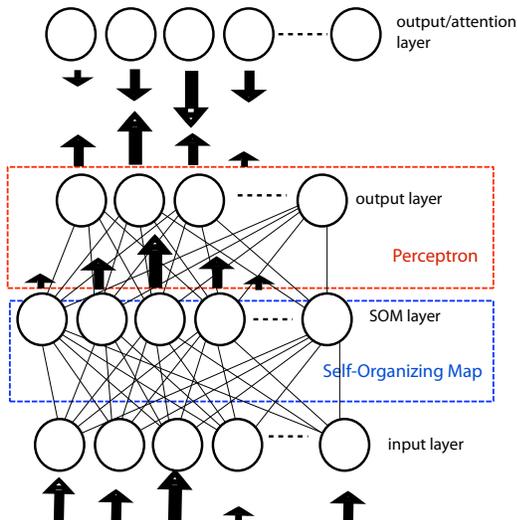


Fig. 1. Architecture of the model investigated in this paper.

The model architecture investigated in this paper is illustrated in Figure 1. It consists of a layer of input nodes, a SOM layer, and a perceptron layer. Input nodes represent the feature decomposition of an input stimulus. We use

a one dimensional feature space in this paper to simplify the illustration, but a generalization to higher dimensions is straight forward. A feature value,  $x_0$ , of an input stimulus is represented by a Gaussian activity packet with mean  $x_0$  in the input layer. Representations in the SOM layer are self-organized by organizing the connections from input to SOM layer with the formalism explained above. The connections between the SOM layer and output layer are learned with supervised learning (delta rule). This supervised learning is intended to simplify a more general reinforcement learning scheme. The signal used for supervision is also a Gaussian activity pattern. Finally, for some of the experiments discussed in this paper, we included some top-down attentional mechanisms. These attentional effects are modeled here by multiplying the input signals with a Gaussian function centered around a chosen signal component that corresponds to the focus of attention in each trial.

### III. RESULTS

We first examine the abilities of the SOM to self-organize when presented with multiple inputs. The network was initialized with uniformly distributed random weights between the input layer and the cortical SOM layer. We presented input patterns consisting of the superposition of ten Gaussians at random locations. An example input is shown on the left in Figure 2A, while the weight matrix after learning on 1000 patterns is shown on the right. The weights show systematically organized receptive fields, demonstrating that the SOM was able to self-organize. While the quality of the organization is less than organizations reported later, we verified that the quality is sufficient to roughly maintain an activity package in the network around the location where it was initialized. Learning with multiple inputs works because the SOM functions as a WTA network and picks one large peak that corresponds to two objects with a similar feature. That is, saliency will focus the attention of the network to one dominating feature, and learning will concentrate on this feature for this particular trial since the postsynaptic activity is suppressed for dissimilar features.

While the network can somewhat organize with multiple inputs, the Hebbian learning of the multiple inputs does introduce considerable noise. It is therefore possible that additional mechanisms are used in natural systems. In particular, there is some evidence that top-down mechanisms are important, including attention [25] and goal directed behavior [6]. To simulate an attentional bias in our model, we chose one of the peaks and modulated the input around this feature with a Gaussian attentional window. The results, shown in Figure 2B, confirm this assertion. In addition, noise in the data can further improve the results as shown in Figure 2C.

While the SOM is able to organize without top-down attentional filtering of multiple inputs, attention is essential for supervised learning based on the SOM representation. To demonstrate this, we included a perceptron on top of the SOM and trained it on inputs with 10 Gaussians at random locations (feature values) after training the SOM. However, only one of these Gaussians was used as supervisor signal.

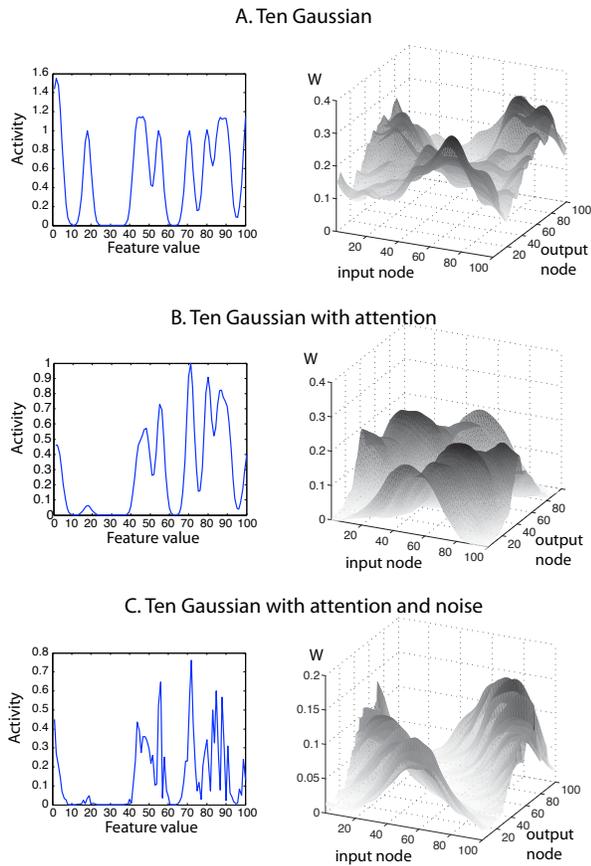


Fig. 2. Demonstration of self-organization with ten randomly placed Gaussian inputs (upper row). The resulting SOM weights are shown in the lower row. Note that the example in the middle self-organized in a rotated direction from the other two examples.

The perceptron could not learn under these conditions. The reason for this is that the SOM WTA network would pick one of the Gaussians randomly. However, with some attentional bias, the SOM WTA networks easily pick the right feature to be learned in the perceptron (Figure 3).

Recent experimental findings by Zhou and Merzenich [6] have shown that the auditory cortex can develop tonotopic maps in adulthood even when the cortex did not organize properly during a critical period in infancy. Crucial for this late learning seems some goal-directed training, which indicates the importance of top-down aspects in self-organizing maps. We have previously modeled the ability to self-organize in adulthood with top-down control with Kohonen’s model [26]. However, a crucial finding of Zhou and Merzenich [6] was that it was even difficult to find preferred stimuli in the experimental conditions that prevented self-organization of the auditory map.

The development of tuning curves in the basic physiological model is shown in Figure 4. Before learning, neurons respond randomly to input features (Figure 4A). After some learning, nodes respond to a cluster of features in some interval, but their responses are still unevenly distributed (Figure 4B). After some further training, smooth tuning

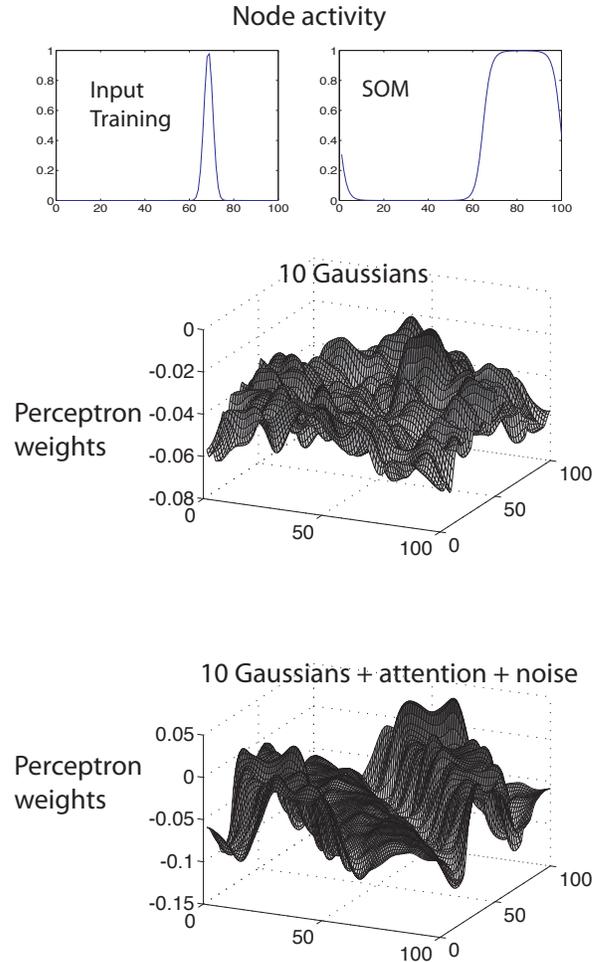


Fig. 3. Demonstration of perceptron learning with and without attention.

curves develop (Figure 4C). The development of the tuning curves is quite robust, so that the lack of clear tuning curves during development is a clear indication of a failed self-organization.

#### IV. CONCLUSION AND OUTLOOK

We have demonstrated that the basic physiological SOM model can self-organize with multiple inputs, but we also demonstrated that attentional biases can considerably improve the organization. This is similar to the benefits of anatomical mini-patches in [14]. Furthermore, attention, such as signal filtering resulting from learning goal-directed behavior is essential to use SOM representations in supervised mapping networks. This is consistent with the findings of Zhou and Merzenich [6]. We also demonstrated the development of tuning curves in such models which was missing from our previous study [26].

There are several additions that we are currently investigating. While the attentional effect was simply simulated by direct input modulation, topographic connections through the system should be used for a more realistic investigation. Furthermore, the learning system on top of the SOM

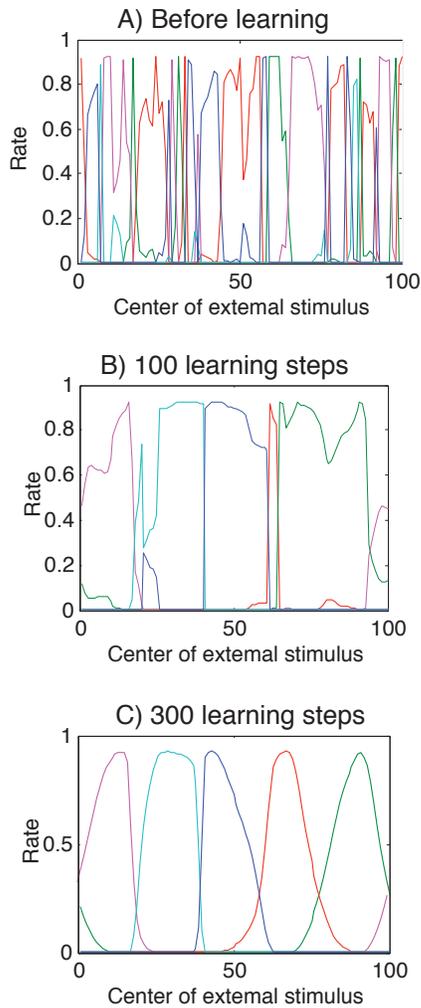


Fig. 4. Development of tuning curves in the basic physiological SOM model.

was a simple perceptron. This should be replaced with a reinforcement learning system to simulate more closely the experimental conditions. Our goal for the development of these more biological models is the investigation of different rehabilitation strategies after lesioning parts of the network.

#### ACKNOWLEDGMENT

This research was supported by NSERC and the FUN international exchange program. TT would like to thank Teuvo Kohonen for some useful discussions.

#### REFERENCES

- [1] D. Hubel and T. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *Journal of Physiology, London*, vol. 160, pp. 106–154, 1962.
- [2] G. Romani, S. Williamson, and L. Kaufman, "Tonotopic organization of the human auditory cortex," *Science*, vol. 216, pp. 1339–1340, 1982.
- [3] J. Kaas, R. Nelson, M. Sur, C. Lin, and M. Merzenich, "Multiple representations of the body within the primary somatosensory cortex of primates," *Science*, vol. 204, pp. 521–523, 1979.

- [4] D. Rasmuson, "Reorganization of raccoon somatosensory cortex following removal of the fifth digit," *J. Comp. Neurol.*, vol. 205, pp. 313–326, 1982.
- [5] C. R. Almlil and S. Finger, "Neural insult and critical period concepts," in *Sensitive Periods in Development: Interdisciplinary Perspectives*, M. H. Bornstein, Ed. Lawrence Erlbaum, Hillsdale, NJ, 1987, p. 123143.
- [6] X. Zhou and M. M. Merzenich, "Intensive training in adults refines all representations degraded in an early postnatal critical period," *Proc. Natl. Acad. Sci.*, vol. 104, pp. 4423–4428, 2007.
- [7] C. von der Malsburg, "Self-organization of orientation sensitive cells in the striate cortex," *Kybernetik*, vol. 14, pp. 85–100, 1973.
- [8] D. J. Willshaw and C. von der Malsburg, "How patterned neural connexions can be set up by self-organisation," *Proc Roy Soc B*, vol. 194, pp. 431–445, 1976.
- [9] S. Amari, "Dynamics of pattern formation in lateral-inhibition type neural fields," *Biological Cybernetics*, vol. 27, pp. 77–87, 1977.
- [10] S. Amari and A. Takeuchi, "Mathematical theory on formation of category detecting nerve cells," *Biological Cybernetics*, vol. 29, pp. 127–136, 1978.
- [11] K. Obermayer, H. Ritter, and K. Schulten, "A principle for the formation of the spatial structure of cortical feature maps," *Proc. Natl. Acad. Sci. USA*, vol. 87, pp. 8345–8349, 1990.
- [12] K. Obermayer, G. Blasdel, and K.J.Schulten, "Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps," *Phys Rev A*, vol. 45, pp. 7568–7589, 1992.
- [13] J. Sirosh and R. Miikkulainen, "Cooperative self-organization of afferent and lateral connections in cortical maps," *Biological Cybernetics*, vol. 71, pp. 66–78, 1994.
- [14] —, "Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex," *Neural Computation*, vol. 9, pp. 577–594, 1997.
- [15] R. Miikkulainen, J. A. Bednar, Y. Choe, and J. Sirosh., *Computational Maps in the Visual Cortex*. Springer, 2005.
- [16] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, pp. 59–69, 1982.
- [17] M. Polla, T. Honkela, and T. Kohonen, "Bibliography of self-organizing map (som) papers: 2002-2005 addendum," *Neural Computing Surveys*, 2007.
- [18] M. Oja, S. Kaski, and T. Kohonen, "Bibliography of self-organizing map (som) papers: 1998-2001 addendum," *Neural Computing Surveys*, vol. 3, pp. 1–156, 2003.
- [19] P. Andras, "Kernel-kohonen networks," *International Journal of Neural Systems*, vol. 12, pp. 117–135, 2002.
- [20] D. M. Donald and C. Fyfe, "The kernel self organising map," in *Proceedings of 4th International Conference on knowledge-based intelligence engineering systems and applied technologies*, 2000, pp. 317–320.
- [21] M. Martn-Merino and A. Muoz, "Extending the som algorithm to non-euclidean distances via the kernel trick," in *Neural Information Processing*, vol. 3316. Springer, 2004.
- [22] T. Kohonen, "Self-organizing neural projections," *Neural Networks*, vol. 19, pp. 723–733, 2006.
- [23] S. Stringer, T. Trappenberg, E. Rolls, and I. Araujo, "Self-organising continuous attractor networks and path integration: One-dimensional models of head direction cells," *Network: Computation in Neural Systems*, vol. 13, pp. 217–242, 2002.
- [24] T. P. Trappenberg, *Fundamentals of computational neuroscience, second edition*. Oxford University Press, 2010.
- [25] N. Suga, J. Yan, and Y. Zhang, "Cortical maps for hearing and ego-centric selection for self-organization," *Trends in Cognitive Science*, vol. 1, pp. 13–20, 1997.
- [26] T. Trappenberg, P. Hartono, and D. Rasmuson, "Top-down control of learning in biological self-organizing maps," in *Lecture Notes in Computer Science, WSOM 2009*, J. Principe and R. Miikkulainen, Eds., vol. 5629. Springer, 2009, pp. 316–324.