



Self organized mapping of data clusters to neuron groups

Dieter Müller*

Leibniz Universität Hannover, Institut für Praktische Informatik, Welfengarten 1 D 30167 Hannover, Germany

ARTICLE INFO

Article history:

Received 2 November 2006

Received in revised form 2 December 2007

Accepted 22 September 2008

Keywords:

SOM
SOM variants
Self organization
Data clusters
Neuron groups
Time dependency
Adaptation procedure
Cortical mapping
Focal dystonia

ABSTRACT

T. Kohonen's self organizing map (SOM) may be considered as a plausible structure for modelling pattern recognition processes in the brain. Neighborhood preservation corresponds closely to what is called somatotopy in the neurosciences, and the context specificity of mappings observed (e. g. in malfunctions of the brain) becomes easily explicable in the framework of the SOM. However, there are two features which impair the aptitude of the classical SOM for neurophysiological models: the adaptation procedure is explicitly time dependent and the procedure consumes the whole set of disposable neurons. Because of the latter property, a SOM cannot learn different tasks, adapting one subset of neurons to a data set X^1 and another to a subsequently presented data set X^2 .

The present paper describes a modified SOM which avoids the drawbacks mentioned above. Its adaptation procedure is time independent. When the training sequence consists of data from successive data clusters X^k each cluster is mapped to a subset G^k of the neuron set G while the other neurons are left almost unchanged. The behavior of the resulting DCNG-SOM is demonstrated in several experiments.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Kohonen's famous paper on self-organized maps (Kohonen, 1982, 1995) has given rise to an enormous number of publications. The original Kohonen map is inspired from the observation that, in the mammalian brain, the mapping of sensory inputs to the cortex is somatotopic (i.e. that signals from neighboring body areas are mapped to neighboring cortex regions). This concept has proven very successful in practical applications. However, it is based on intuitive considerations, and not on a strict mathematical theory. Therefore, it suffers from certain deficiencies which are discussed in Kohonen (1995).

The original SOM and most of its variants require that the learning parameters (i.e. learning strength and range of adaptation within the neuron grid) are successively modified during the training phase. It is necessary to define a so called annealing scheme, controlling this time dependency. Unfortunately there is no firm theoretical basis for constructing such schemes, and often they are determined empirically. That the mapping changes with time is the purpose of the process. However the time dependency is an explicit one: learning parameters vary with time and thus the quantitative properties of the process itself change during the training phase. This prevents the classical SOM from learning a new mapping when the training is complete.

Meanwhile, there are a large number of papers which define new SOM types. Most of them have a mathematical basis and some of them avoid the explicit time dependency.

Typical examples are the SOAN (self organization with adaptive neighborhood neural network) (Iglesias & Barro, 1999) and the parameterless PLSOM (Berglund & Sitte, 2006). Both use the current mapping error to adjust the internal parameters of the adaptation process. In the time-adaptive SOM (TASOM), (Shah-Hosseini & Safabakhsh, 2003) each node has its own variable learning strength and neighborhood size. As a consequence, the network as a whole keeps its ability to learn independent from time. Somewhat simplified, the TASOM approach may be considered as shifting the dependency on externally controlled learning parameters to a dependency on individual node properties, thus converting the explicit time dependency into an implicit one. The present paper pursues a similar strategy. A rather elaborate mathematical procedure is used in the auto-SOM (Haese, 1999; Haese & Goodhill, 2001): here the weight vectors are adapted by means of a Kalman filter and the learning parameters are determined so as to minimize the prediction error variance. With regard to the result, the generative topographic mapping (GTM Bishop, Svensén, and Williams (1998)) is also a SOM. The adaptation algorithm is probabilistic and does not need a decreasing learning strength or a shrinking range of adaptation. However, the mapping problem is formulated in terms of a latent variable model, and thus the connection to Kohonen's approach is weak. In some SOM variants, the number of neurons and the whole network structure change with time. Although this is a very strong time dependency, it is an implicit one: addition and

* Tel.: +49 05723 81221, +49 0511 762 3677(O); fax: +49 0511 762 19679.
E-mail address: dm@inf.uni-hannover.de.

removal of neurons are controlled solely by the stream of input signals. Examples are the growing cell structure (Fritzke, 1994a), the growing neural gas (Fritzke, 1995, 1994b) and the plastic self-organizing map (Lang & Warwick, 2002).

Compared to the original Kohonen SOM, these variants are remarkable improvements. They do not need an external control of parameters and, especially the TASOM, is well suited for tasks with changing data sets. Above all, they achieve a better mapping quality, at least in certain applications

On the other hand, mapping quality and performance become less important if SOM structures are used to model processes in the brain. Here we are confronted not only with excellent function, but also with malfunction occurring in certain situations. Therefore it seems useful to study another kind of SOM which is designed primarily for biological modelling purposes, and not for better performance.

Soon after the first SOM publications, Merzenich et al. demonstrated that the reorganization observed in the sensory cortex of mammals after peripheral nerve damage can be modeled as an adaptation process in a Kohonen map (Merzenich et al., 1984). Martinetz, Ritter, and Schulten (1988) showed, that the auditive cortex of a bat may be considered as a neighborhood preserving map of the relevant space of ultrasonic signals. Meanwhile the SOM concept is accepted as a tool for understanding the processing of sensory signals in the brain (Kaas, 1991; Sirosh & Miiikulainen, 1995; Turrigiano & Nelson, 2004; Wiemer, Spengler, Joubin, Stagge, & Wacquant, 2000). Nevertheless, at least the classical SOM has two properties which hamper its use for the modelling task described above:

1. An explicit time dependence means that there exists a control mechanism outside the SOM. When the training for a task is completed, this mechanism would be responsible for restoring the plasticity if a new task is to be learned. The time independent SOM variants mentioned above, avoid this problem, but their mathematical procedures cannot easily be transferred into a biological framework.
2. The training process of the classical SOM adapts the whole set of disposable neurons to a given set of signal vectors. There remain no neurons which can be used for a subsequent training process with a different set of signal vectors. This is in contradiction to the observed plasticity of the brain. We can learn many new tasks without forgetting those we have learned before.

After all, it appears useful to modify the Kohonen training procedure, retaining the principal properties of the original SOM and avoiding the undesirable properties pointed out above. The necessary modifications are the following:

- The adaptation procedure must not contain an explicit time dependency.
- Training with a certain data set X^1 should result in a mapping of X^1 to a certain neuron group $G^1 \subset G$ and not to the whole neuron set G . The bulk of neurons should be left disposable for subsequent training phases with different data sets X^1, X^2, \dots

The SOM introduced in the present paper fulfills these requirements. However in contrast to the time independent SOMs mentioned above, it is constructed in a purely intuitive way. The goal is not an improved mapping, but a neurophysiologically plausible mechanism. Rather than from a mathematical theory the design starts from the idea, that a firing neuron remains for some time in an excited state with increased sensibility and that its excitation spreads to his neighbors. As in the TASOM, a particular neuron has not only its own synaptic coupling vector \vec{w} but, in addition, a state which determines its behavior. The state, however, is only a single entity a , the activation. This is enough to allow for a time independent adaptation procedure which does not use all

available neurons, but only a certain neuron group, to represent a data cluster. Because of this property, the resulting SOM is called DCNG-SOM. As in the growing neuron structures (Fritzke, 1994a, 1995; Lang & Warwick, 2002), the incoming stream of data controls the number of neurons involved in the mapping of a cluster. In contrast to these SOM variants, the neurons are not taken from a potentially infinite stock. Instead, the number of neurons is fixed, and they have a fixed position in physical space.

The aspect of performance in applications is not considered. We simply study the properties of the DCNG-SOM resulting from the use of the modified neuron. This is because the long-term objective is not only to model neuroplasticity, but also to explain disorderings in the mapping of sensory signals to the sensory cortex. Such disorderings have been observed (e. g. in patients suffering from focal dystonia).

Focal dystonia is a movement disorder occurring in several forms, as for example, writers cramp and musicians cramp. Apparently the cause is not an organic defect but rather something like an overtraining. This suggests that focal dystonia has to do with a misled self organization process. Sanger and Merzenich (2000) have hypothesized that it is a manifestation of an unstable sensorimotor control loop. They discuss several mechanisms which possibly could lead to a gain > 1 in the feedback loop. The mapping mentioned above is a link in this loop. As yet, theoretical studies on focal dystonia concentrate on the time behavior of sensory signals and on the control theoretic aspects of the problem. However, in some musicians with focal dystonia, the disordered cortical representation of digits could be observed directly by functional magnetic resonance imaging (Elbert et al., 1998). The increasing empirical material in this field suggests the study of the formation of disordered mappings with SOM based models

It is a characteristic of focal dystonias that they are task specific. The focal dystonia of musicians affects the control of finger movements only in the context of instrument playing. It does not affect the function of the same fingers in other activities. This means that the disturbed mapping is effective only in a special context. The context specificity becomes understandable if a cortical region is described as a SOM. A sensory stimulus might be considered as a signal vector

$$\vec{x} = (x_1, \dots, x_n) = (s_1, \dots, s_k, c_1, \dots, c_{n-k}) = (\vec{s}, \vec{c}) \quad (1)$$

where \vec{c} describes the context. In this picture, it appears plausible that certain stimuli $(\vec{s}^{(i)}, \vec{c})$ are mapped to a corrupted part of the map, while stimuli with the same $\vec{s}^{(i)}$ and a different context part \vec{c}' are processed in correctly working regions. Admittedly, it remains an open question as to whether sensory stimuli are actually encoded as high dimensional signal vectors.

In the literature, there are some context-aware SOM variants (e.g. the recursive self-organizing map by Voegtlin (2002)). They are designed to represent the temporal context of patterns, and in principle they allow the reconstruction of pattern sequences. In contrast, the present paper restricts itself to the more simple problem of constructing different maps for data clusters which are presented successively to the net.

2. Mapping of data clusters to neuron Groups: The training procedure

Let us assume a data set

$$X = X^1 \cup X^2 \cup X^3 \dots \cup X^k \subset R^m \quad (2)$$

with the property

$$\exists M \in (0, 1.0) : \forall a, b \in X^s, u, v \in X^t : |a - b|, |u - v| < M|a - u|, \quad s \neq t \quad (3)$$

i.e. X consists of distinct clusters X^s .

Download English Version:

<https://daneshyari.com/en/article/404604>

Download Persian Version:

<https://daneshyari.com/article/404604>

[Daneshyari.com](https://daneshyari.com)