Sleep Spindle Detection by Using Merge Neural Gas

Pablo A. Estévez, Ricardo Zilleruelo-Ramos, Rodrigo Hernández, Leonardo Causa and Claudio M. Held
Department of Electrical Engineering, Universidad de Chile
Casilla 412-3, Santiago, Chile
email: {pestevez, rzilleru, rodrhern, lcausa, cheld}@ing.uchile.cl

Keywords: self-organizing maps, neural gas, sleep spindles, EEG

***Abstract*— In this paper the Merge Neural Gas (MNG) model is applied to detect sleep spindles in EEG. Features are extracted from windows of the EEG by using short time Fourier transform. The total power spectrum is computed in five frequency bands and used as input to the MNG network. The results show that MNG outperforms simple neural gas in correctly detecting sleep spindles. In addition the temporal quantization results as well as sleep trajectories are visualized on two-dimensional maps by using the OVING projection method.

1 Introduction

Self-organizing neural networks [10] have been recently extended for processing sequential and tree-structured data [2][8][15][17]. Hammer et al. [7] presented a general framework of recurrent self-organizing network models, based on the context definition, i.e. the way how sequences are internally represented.

In the Merge SOM (MSOM) [15] approach, the current input is considered along with the context of the previous time step. The context is compactly described by a linear combination (merge) of the weight and the context of the last winner neuron. It has been proved that MSOM context leads to an efficient fractal encoding as the fixed point of the training dynamic. MSOM can simulate finite automata thanks to its explicit representation of the context as an independent part of the neurons. Since MSOM context does not depend on the lattice architecture, it can be combined with other self-organizing neural networks such as Neural Gas (NG) [12]. The resulting model is called Merge Neural Gas (MNG) [16]. The SOM vector quantization is constrained by a predefined lattice or output grid. The NG algorithm overcomes this constraint by not defining an output space. As a consequence, the neural gas is in general a better quantizer than SOM, outperforming the latter when quantizing topologically arbitrary structured manifolds [12]. However, the lack of an output space has restricted the applications of neural gas to visualization. Recently, in [3] an output space was defined for the neural gas network for visualization purposes. This allows projecting the neural gas quantization results (codebook vectors) from multidimensional spaces onto two-dimensional maps.

Human sleep is usually divided in five stages, one stage called rapid eye movement (REM) sleep, and four stages of non-REM sleep [9]. Encephalographic (EEG) signals are analyzed to establish the presence or absence of sleep spindles (SSs), which are sequences (trains) of fast (sigma) EEG waves, lasting more than 0.5 s and with a magnitude above 10 μV [4]. The frequency of sleep spindles ranges between 10 and 15 Hz. SSs are normally detected in the anterior EEG derivations (FP-C channels) during non-REM sleep stage II. Fig. 1 shows an example of SS activity in EEG. Usually over one thousand SS events occur during one night recording, so an automatic recognition procedure is highly desirable. Many methods for SS detection have been proposed in the literature. In [1] a two stage classification procedure is proposed. Features are extracted from windows of the EEG by using an autoregressive model. The first stage eliminates definite non-SS patterns by using a discrete perceptron. In the second stage a support vector machine (SVM) classifier is used. The results reported are an average sensitivity of 94.6% and an average false detection rate of 4.0% on six recordings for testing. In [6] the short time Fourier transform is used for feature extraction. The EEG signal is first multiplied by a Hamming window and then the Fourier transform of the windowed signal is taken. The feature vector contained 32 coefficients between 2 and 64 Hz. This feature vector was used as input to an MLP classifier (88.7% accuracy) and an SVM classifier (95.4% accuracy). Another approach is to use wavelet transform [14].

In this paper an application of the MNG model to sleep spindle detection in EEG is presented. We focus here on

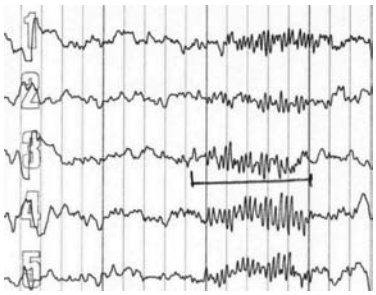

Figure 1: **Example of SS activity in five EEG channels.**

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany, ISBN 978-3-00-022473-7
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

the neural gas model since it is a better quantizer than SOM and its results can be projected onto a two-dimensional map using an adaptation of the OVING projection method [3]. The MNG model is compared with NG on the task of sleep spindle detection in real EEG recordings.

2 Merge Neural Gas

In this section a brief overview of MNG is given. Neurons are a vector tuple $\left(\mathbf{w}^i, \mathbf{c}^i\right) \in \Re^d \times \Re^d$, where d correspond to the dimensionality of the input signals.

Given the current entry $x(n)$ of a sequence, the best matching neuron I_n is the closest neuron according to the following recursive distance criterion:

$$d_i(n) = (1-\alpha) \cdot ||x(n) - w^i||^2 + \alpha \cdot ||c(n) - c^i||^2 \quad (1)$$

where the current context, $c(n)$, corresponds to a linear combination of the weight and context of the previous winner, I_{n-1}, i.e. the best matching unit in the last time step. The current context is defined as:

$$c(n) = (1-\beta) \cdot w^{I_{n-1}} + \beta \cdot c^{I_{n-1}}. \quad (2)$$

Training takes place by adapting both weight and context towards the current input and its context representation, respectively. Adaptation is made by Hebbian learning as shown in the following equation:

$$\begin{aligned} \triangle w^i &= \epsilon_w(k) \cdot h_{\lambda(k)}\left(r(d_i, \mathbf{d})\right) \cdot \left(x(n) - w^i\right) \quad (3) \\ \triangle c^i &= \epsilon_c(k) \cdot h_{\lambda(k)}\left(r(d_i, \mathbf{d})\right) \cdot \left(c(n) - c^i\right) \end{aligned}$$

where k is the current training epoch, d is a vector whose components are all the neuron distances at time n, and $r(d_i, d)$ denotes the rank of the ith neuron with respect to the current input and its context representation. $h_{\lambda(k)}\left(r(d_i, d)\right)$ represents the neighborhood ranking function which is defined as,

$$h_{\lambda(k)}\left(r(d_i, \mathbf{d})\right) = exp\left(\frac{-r(d_i, d)}{\lambda(k)}\right) \quad (4)$$

where $\lambda(k)$ controls the width of the neighborhood function and is defined as follows,

$$\lambda(k) = \lambda_0 \cdot \left(\frac{\lambda_f}{\lambda_0}\right)^{\frac{k}{k_{max}}} \quad (5)$$

where k_{max} is the maximum number of epochs. The learning rates $\epsilon_w(k)$ and $\epsilon_c(k)$ are functions of the same type as (5). The current context $c(n)$ is updated during training in the way described by equation (2).

It has been proved that Hebbian learning converges to the following optimal weight and context vectors at time n (i.e., are stable fixed points):

$$\begin{aligned} w^{opt(n)} &= x(n) \quad (6) \\ c^{opt(n)} &= \sum_{j=1}^{n-1} (1-\beta) \cdot \beta^{j-1} \cdot x(n-j), \end{aligned}$$

provided that there are enough neurons and neighborhood cooperation is neglected, i.e. for late stages of learning [7].

3 OVING Projection Algorithm

In order to project multidimensional vectors onto two-dimensional maps we adapted a projection method called OVING [3], which provides an output representation to Neural Gas. In the original formulation, the projection part is done simultaneously with the vector quantization. Here only the projection part of OVING is applied to the result of the temporal vector quantization of MNG. The initial topology of the network is a set of m neurons. Each neuron j has associated a $2 \cdot N$-dimensional codebook vector, $[w^j|c^j]^T$ in input space and a two-dimensional codebook position in output space, z_j, for $(j = 1, \ldots, m)$. The following global cost function is minimized:

$$E = \frac{1}{2} \sum_{j=1}^{N} \sum_{k \neq j} (D_{j,k} - d_{j,k})^2 F(s_{j,k}) = \frac{1}{2} \sum_{j=1}^{N} \sum_{k \neq j} E_{j,k}, \quad (7)$$

where

$$d_{j,k} = (1-\alpha) \cdot ||w^j - w^k|| + \alpha \cdot ||c^j - c^k|| \quad (8)$$

is the distance between the jth and kth codebook vectors in input space, and

$$D_{j,k} = ||z_j - z_k|| \quad (9)$$

is the distance between the jth and kth codebook positions in output space. The function F is defined as

$$F(f) = e^{-\left(\frac{f}{\sigma(t)}\right)} \quad (10)$$

where $\sigma(t)$ is the width of the neighborhood that decreases with the number of iterations as in eq. (5).

(1) Initialize the codebook positions z_j randomly.
(2) Present a codebook vector $[w^{j^*}, c^{j^*}]^T$ $(j^* = 1, \ldots, m)$.
(3) Generate the ranking in output space $s_{j^*j} = s(z_j^*(t), z_j(t)) \in \{0, 1, \ldots, m-1\}$ for each vector position $z_j(t)$ with respect to the vector position $z_j^*(t)$, corresponding to the projection of codebook vector $[w^{j^*}, c^{j^*}]^T$.
(4) Update codebook positions:

$$\begin{aligned} z_j(t+1) &= z_j(t) + \gamma(t) F(s_{j,j^*}) \\ &\times \frac{(D_{j,j^*} - d_{j,j^*})}{D_{j,j^*}} (z_{j^*}(t) - z_j(t)) \quad (11) \end{aligned}$$

where $\gamma(t)$ is the learning rate, which decreases with the number of iterations t, in the same form as eq. (5).
(5) If $t < t_{max}$ go back to step 2.

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany, ISBN 978-3-00-022473-7
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

The OVING projection algorithm tends to preserve well the local neighborhood [3]. In contrast Sammon's mapping [13] preserves globally the interpoint distances between the input and output spaces. We present below projections using both methods and compare their performance using the topology preservation measurement q_m [11].

4 Methods

Two 18-channel polysomnographic recordings of different healthy infants, sampled at 250 Hz were used. These recordings include five EEG channels and other signals such as EOG, EKG, muscle tone, limb movements, etc. For each recording, a continuous sample of 45 minutes was drawn from EEG channel 1 containing all non-REM sleep stages. One recording was used for training and the another for testing. The EEG signal was pre-processed with a Hamming window of 128 samples (0.512 s). Then the fast Fourier transform (FFT) was taken for the windowed signal. A step size of 16 samples (0.064 seconds) between windows was used, i.e. windows are overlapped. The FFT is computed as

$$X[k] = \sum_{n=0}^{N-1} x(n) W_N^{nk} \quad k = 0, 1, \cdots, N-1 \qquad (12)$$

where $x[n]$ is the windowed EEG data, N is the number of samples in a window, and $W_N = e^{-j\frac{2\pi}{N}}$. The power spectrum for each coefficient is computed as

$$P[k] = \frac{X[k] \times X[k]^*}{N} \qquad (13)$$

where $X[k]^*$ is the complex conjugate of coefficient $X[k]$. The total power in the following five frequency bands was computed: delta-theta [0.5,7] Hz, alfa [7,10] Hz, sigma [10,15] Hz, beta [12,30] Hz and high-frequency band [30,60] Hz. The total power for each band is computed as follows:

$$P_{band}[a, b] = \sum_{k \geq a}^{k \leq b} P[k] \qquad (14)$$

where a is the lower frequency of the band, and b is the higher frequency of the band. A human expert labeled each window as containing sleep spindles or not. If only part of the window contained SSs then a number in [0,1] was assigned.

The simulation parameters of NG and MNG were set as follows:
For all simulations, the parameters of the neighborhood function $\lambda(k)$ in (5) were set to initial value $\lambda_0 = m$, i.e. equal to the number of neurons, and final value $\lambda_f = 0.001$. The number of training epochs was set to $k_{max} = 2000$, but in MNG training we use $k_{max} = 1000$ for the first training phase and $k_{max} = 1000$ for the second training phase (see [5]).
For NG the number of neurons was set to $m = 300$. The initial and final learning rate were set to $\epsilon_0 = 0.3$ and $\epsilon_f = 0.01$ respectively.
For MNG first training phase, the number of neurons was set to $m = 30$. The initial and final learning rate, for both weight and context adaptations were set in the same way as follows: $\epsilon_{w_0} = \epsilon_{c_0} = 0.3$ and $\epsilon_{w_f} = \epsilon_{c_f} = 0.01$. The context and merging parameters were set as $\alpha = \beta = 0.0$, thus not taking into account the context influence.
For MNG second training phase, ten replicas were made for every neuron trained during the first phase. It is assumed that these neurons are well located, so the initial neighborhood size is set to $\lambda_0 = 10$, i.e. the number of replicas of each neuron. The weight learning rate was set to $\epsilon_w(k) = 0.01$, and the context learning rate was set as in the first training phase. The context influence was set to $\alpha = 0.45$. The parameter β was selected by making simulations for each value in the interval $[0.0, 0.9]$ with a step of 0.1. The parameter β was chosen as the value that minimized the global classification error in the training set.

To evaluate the quality of the clustering done by the algorithms, a classification task was performed, although MNG and NG use unsupervised training. The network's neurons were associated to a probability of belonging to each of the two classes. A histogram of two bins was computed for each neuron. The EEG signal was swept by shifting the window, and the best matching unit was calculated step by step. The histogram bin of the winner neuron corresponding to the class of the current input was incremented by one. Histograms were accumulated during the swept of the signal. Each neuron was associated to the category having maximum histogram bin value.

After labeling the neurons, the classification task is done step by step. Each window is assigned to the category of the best matching unit. This was done for both the training and the testing data sets. The network classification is contrasted with the labels given by an expert.

5 Results

Table 1 shows the results obtained on the training set for $\beta = 0.4$, value for which the minimum classification error was obtained. The true positives (TPs) are defined as windows containing SSs. False positives correspond to windows that do not contain SSs but are labeled as containing SSs by the model. The accuracy is the global classification rate using two categories. Table 2 shows the results obtained on the testing set for $\beta = 0.4$. It can be observed that the MNG model obtained a higher true positive rate than NG maintaining the false positive rate. It achieved a higher global accuracy too, showing the influence of taking into account the context. Tables 1 and 2 show the classification results per window of fixed size. The real SS events however have a variable duration between 0.5 s and 3 s. Table 3 shows the classification results for SS events. An

SS event is considered detected if it is detected partially or completely. The results show that nearly 90% of the SS events were correctly detected in the testing set. The expert agreement indicates what percentage of the total number of events detected by the MNG model was labeled as SS event by the human expert. The results show that nearly 63% of the events detected by the MNG model are real SS events.

Fig. 2 shows an example of true positive detection of an SS event. Fig. 3 shows an example of false negative detection of an SS event. It can be seen that both the left and right borders of the SS event are not detected by the MNG model. The MNG model shows some difficulty in getting the precise location of the beginning and ending of the SS events.

Fig. 4 shows the projection of the neurons obtained with MNG onto a two-dimensional map by using OVING. Fig. 5 shows the projection of the same data by using Sammon's mapping. In both maps it can be observed a good separation of the SS and non-SS categories.

Table 1: SS classification results on the training set using overlapped EEG windows

	Training		
Model	TP[%]	FP[%]	Accuracy[%]
NG	70.1%	1.3%	92.5%
MNG	74.7%	1.5%	96.4%

Table 2: SS classification results on the testing set using overlapped EEG windows

	Testing		
Model	TP[%]	FP[%]	Accuracy[%]
NG	52.3%	0.7%	95.8%
MNG	61.5%	0.8%	96.3%

Table 3: Classification results of SS events on the training and testing sets using the MNG model

Recording	Detection rate [%]	Expert agreement [%]
Training	96.6	71.8
Testing	89.7	62.9

Table 4: Topology preservation measurement for OVING and Sammon mapping

Method	NG	MNG
OVING	0.458	0.341
Sammon	0.444	0.303

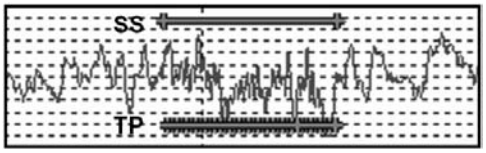

Figure 2: **Example of SS true positive detection.**

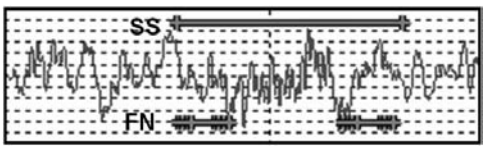

Figure 3: **Example of false negative detection at the left and right borders of the SS event.**

Table 4 shows a comparison of both projections using the topology preservation measurement q_m, which measures the quality of the local neighborhood preservation. The results show that OVING obtained a higher q_m than Sammon's mapping. Fig. 6 shows a trajectory of EEG windows within the map that starts in the SS region and finishes in the non-SS region.

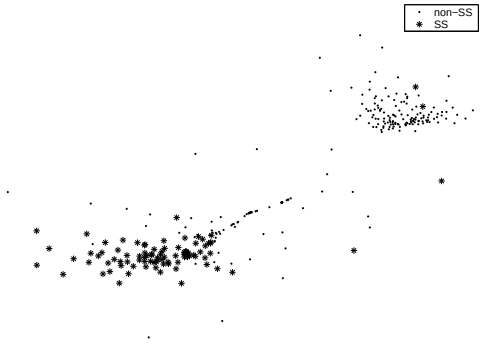

Figure 4: **Mapping of MNG neurons by using OVING.**

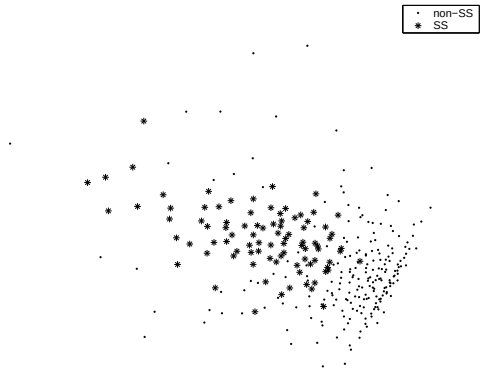

Figure 5: **Projection of MNG neurons using Sammon mapping.**

Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM 2007)
Published by the Neuroinformatics Group, Bielefeld University, Germany, ISBN 978-3-00-022473-7
All contributions to WSOM 2007 are available online at: http://biecoll.ub.uni-bielefeld.de

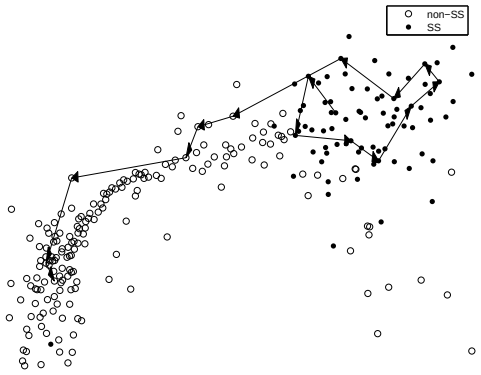

Figure 6: **Map showing the trajectory of part of the EEG, starting from a SS region (.) and ending in a region on non-SS (o).**

6 Conclusions

The MNG model has been applied to the real-world problem of detecting sleep spindles in EEG signals. In this task, the MNG model outperformed NG showing that the context is useful for distinguishing between different temporal sequences. Although MNG is an unsupervised model, the detection rates obtained are surprisingly high. These promising results should be confirmed by using a larger database. As the results may depend on the quality of the feature extraction, other techniques may be implemented and compared to short-time Fourier transform. The MNG model may be used as a pre-processor, in order to indicate which parts of the EEG signals are worth to be analyzed in more detail. The proposed OVING projection method gives to merge neural gas an output representation, which is very useful for visualization tasks.

Acknowledgements

This research was supported by Conicyt-Chile under grant Fondecyt 1050751. The authors would like to thank to Dr. P. Peirano and Mr. M. Garrido from Sleep Lab, INTA, Universidad de Chile, for providing expert analysis of the EEG data.

References

[1] N. Acir and C. Guzelis, "Automatic recognition of sleep spindles in EEG by using artificial neural networks." *Expert Systems with Applications*, vol. 27, pp. 451–458, 2004.

[2] G. J. Chappell and J. G. Taylor, "The temporal Kohönen map." *Neural Networks*, vol. 6, no. 3, pp. 441–445, 1993.

[3] P. A. Estévez and C. J. Figueroa, "Online data visualization using the neural gas network," *Neural Networks*, vol. 19, no. 6, pp. 923–934, 2006.

[4] P. A. Estévez, C.M. Held, C.A. Holzmann, C.A. Perez, J.P. Perez, J. Heiss, M. Garrido, P. Peirano, "Polysomnographic pattern recognition for automated classification of sleep-waking states in infants," *Medical & Biological Engineering & Computing*, vol. 40, pp. 105–113, 2002.

[5] P.A. Estevez, R. Zilleruelo-Ramos, and J.M. Zurada, "Window merge neural gas for processing pattern sequences," *Proceedings of the Int. Joint Conference on Neural Networks (IJCNN)* , Orlando, USA, 2007 (accepted).

[6] D. Gorur, U. Halici, H. Aydin, G. Ongun, F. Ozgen, and K. Leblebiciouglu, "Sleep spindles detection using short time Fourier transform and neural networks," *Proceedings of the Int. Joint Conference on Neural Networks (IJCNN)* , Vol. 2, pp. 1631-1636, 2002.

[7] B. Hammer, A. Micheli, A. Sperduti, M. Strickert, "Recursive self-organizing network models," *Neural Networks*, vol. 17, no. 8-9, pp. 1061–1085, 2004.

[8] B.Hammer, A.Micheli, N.Neubauer, A.Sperduti, M.Strickert, "Self organizing maps for time series", *Proceedings of the Workshop on Self-Organizing Maps (WSOM)*, pp. 115-122, Paris, 2005.

[9] C.M. Held, J. Heiss, P. A. Estévez, C.A. Perez, M. Garrido, C. Algarin, P. Peirano, "Extracting fuzzy rules from polysomnographic recordings for infant sleep classification," *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 1954–1962, 2006.

[10] T. Kohonen, *Self-Organizing Maps*. Berlin, Heidelberg: Springer, 1995.

[11] A. König, "Interactive visualization and analysis of hierarchical neural projections for data mining," *IEEE Transactions on Neural Networks*, V. 11, N 3, pp. 615–624, 2000.

[12] T. M. Martinetz, S. G. Bercovich, and K. J. Schulten, ""Neural-gas" network for vector quantization and its application to time-series prediction," *IEEE Transactions on Neural Networks*, vol. 4, pp. 558–569, 1993.

[13] J.W. Sammon, "A nonlinear mapping for data structure analysis," *IEEE Transactions on Computers*, V. C-18, pp. 401–409, 1969.

[14] S.V. Schonwald, G.J.L. Gerhardt, E.L. de Santa-Helena, and M.L.F. Chaves, "Characteristics of human EEG sleep spindles assessed by Gabor transform." *Physica A*, vol. 327, pp. 180–184, 2003.

[15] M. Strickert and B. Hammer, "Merge som for temporal data." *Neurocomputing*, vol. 64, pp. 39–72, 2005.

[16] M. Strickert and B. Hammer, "Neural gas for sequences," In T. Yamakawa (Ed.), *Proceedings of the Workshop on Self-Organizing Networks (WSOM)* , pp. 53–58, Kyushu, Japan, 2003.

[17] T. Voegtlin, "Recursive self-organizing maps," *Neural Networks*, vol. 15, no. 8-9, pp. 979–991, 2002.