The activation frequency self-organizing map (AFSOM)

Antonio Neme, Pedro Miramontes

Universidad Autónoma de la Ciudad de México

San Lorenzo 290 Col. Del Valle Mexico, D.F. Mexico. neme@nolineal.org.mx

Universidad Nacional Autónoma de México

keywords: self-organization, Hebbian learning, non-radial influence

Abstract— In the self-organizing map (SOM), the best matching units (BMUs) affect neurons as a function of distance and the learning parameter. Here we study the effects in SOM when a new parameter in the learning rule, the activation frequency, is included. This parameter is based on the relative frequency by which each neuron is included in each BMU's neighborhood, so there is an individual memory (synapse strength) of the activation received from each neuron. The parameter leads to non-radial influence areas for BMUs that modifies the map formation dynamics, including the fact that the weight vector for BMU may not be the closest one to the input stimulus after weight adaptation. Two error measures are lower for the maps trained with this model than those obtained with SOM, as shown in experiments with six data sets.

1 Introduction

The self-organizing map (SOM) is presented as a model for the self-organization of neural connections, which is translated in the ability of the algorithm to produce organization from disorder [5]. One of the main properties of the SOM is its ability to preserve in the output map the topographical relations present in input data [16], This property is achieved through a transformation of an incoming signal pattern of arbitrary dimension into a low-dimensional discrete map and to adaptively transform data in a topologically ordered fashion [16]. Each input data is mapped to a BMU, which affects other neurons accordingly to the learning equation:

$$w_n(t+1) = w_n(t) + \alpha_n(t)h_n(g,t)(x_i - w_n(t)) \quad (1)$$

Where $\alpha(t)$ is the learning rate at time t and $h_n(g,t)$ is the neighborhood function from BMU neuron g to neuron n at time t. In general, the neighborhood function decreases monotonically as a function of the distance from neuron g to neuron n. This decreasing property has been stated to be a necessary condition for convergence [7, 6]. The SOM tries to preserve relationships of input data by starting with a large neighborhood and reducing it during the course of training [16]. Neighborhood and learning parameters are reduced by an annealing scheme although the form they take is not critical [16].

As pointed out in [29], SOM follows the idea of using a deformable lattice to transform data similarities into spatial relationships. The lattice is deformed by applying learning equation (1) to the neurons in the network. Here, we propose an additional parameter for equation (1) that quantifies the influence a BMU n has over the neurons in the network as a function of the relative frequency with which n includes them in its neighborhood. It also measures the influence each input vector m has on them as a function of the number of times the BMU for m affects the neurons. This frequency activation parameter allows non-radial neighborhood and, as reported in the results, forms better maps, in terms of two error measures.

Although several modifications have been proposed to the SOM learning rule, they do not reflect, at least to our knowledge, the frequency of activation from other neurons. Some works incorporate non-radial influence from BMUs to neighbors, as, for example, in [19] it is proposed the recursive Fisherman's rule and some hybrid rules that reflect an attenuation of the adaptation as the distance from the BMU to the affected neuron increases. In [12] the activity patterns are non-radial and determined by a mechanism based on a cooperative information control.

One of the first works that incorporated the concept of memory for each neuron was proposed in [4], in which an activation memory is defined, in order to identify the new active neuron, and a modification in the BMU selection mecanism is presented. Also, a SOM-related model has been studied in the light of reaction-diffusion mechanisms where the BMU perturbes the excitable media and generates a symmetrical traveling wave [26]. Several models of the visual cortex have been proposed [23, 21, 1] in which a dynamic Hebbian-like behavior is considered, but there is a radial influence between neurons.

The activity patterns in the cortex are irregular and non-radial [13, 3], that is, neurons in different regions become active for a given stimulus, whereas SOM shows regular activity, since the neighborhood function defines a symmetrical influence area, as well as the learning function, although there are some variations that present a different behavior.

There is biological evidence that connectivity in the brain cortex is not regular and could be approximated by a small-world topology [31, 30], which means that when a given neuron or group of networks become active, they

activate neurons in their proximity as well as those located farther. In [24] it is reported that low error measures are achieved when the topology of the lattice resembles that of a small-world network. So, in self-organizing maps over small-world lattices, non-radial activity patterns which resemble that of the cortex are obtained. Also, in an preliminary study of activation frequency, some non-radial patterns are reported in [25]. In this work, we obtain non-radial activity patterns not by an explicit modification in the lattice topology, but by a modification in the learning process.

As a model of the brain cortex, SOM has been applied in the study of map formation in visual cortex [22, 2], as models of brain maps or as ordered projections from sensorial areas to cortical regions [17]. However, the SOM fails to reproduce the activity patterns mentioned above [18, 9], although some models, as the presented in [20], in which a modification in the SOM adaptation kernel is considered in order to include surround inhibition, achieve the formation of pinwheel patterns similar to those observed in the visual cortex. In these models, however, influence from BMUs is radial and symmetrical.

Here, we propose a modification in the learning equation such that the activity patterns resemble in more detail those non-radial patterns present in the cortex. This non symmetrical influence area formation is achieved through the activation frequency parameter which is a kind of memory between neurons that resembles Hebbian learning. The self-organizing capabilities of the SOM are not limited but enhanced, as shown in section 3.

The proposed modification to the SOM is an additional term in eq. (1) that quantifies the frequency by which a neuron is affected by each BMU. We apply this modification to the bubble neighborhood scheme, as in this scheme it is straightforward to keep track of the number of times each BMU affects others, by including them on its neighborhood.

The variation we propose is the activation frequency, explained in section 2 whereas in section 3 several results are presented. In section 4 the results are discussed and some conclusions are stated.

2 The activation frequency in the SOM (AFSOM)

In Hebbian learning, the strength of the synapsis between two neurons is a function of the frequency by which both of them are fired simultaneously for the same stimulus. Here, we incorporate a Hebbian-like parameter to the SOM learning rule. This parameter leads to the activation frequency SOM, hereafter AFSOM, which allows non-radial influence from BMUs to their neighbors while, at the same time, the maps obtained by it show lower errors than those obtained by SOM.

The activation frequency between neuron k and neuron

n is defined as a function of the number of times k has become a BMU and included neuron n in its neighborhood (and thus affected it), named T_n^k, divided by the number of times n has been affected by any BMU, T_n. $\Omega_n(k)$ is defined as T_n^k/T_n and it represents the influence from neuron k to neuron n.

The activation frequency function $\rho_n(k)$ is a function of $\Omega_n(k)$. Eq. (2) shows the learning equation for the SOM with the new parameter. It is defined in the range $[0, 1]$ so an unconstrained weight growth is avoided, which is a main problem in the original Hebbian learning formulation [8]. When there is no influence between BMU k and neuron n, $\rho_n(k) = 0$. On the other hand, when the influence from neuron k attains a maximum to neuron n, then $\rho_n(k)$ should be 1.

$$w_n(t+1) = w_n(t) + \alpha_n(t)h_n(g,t)\rho_n(k)(x_i - w_n(t)) \quad (2)$$

In order to study the behavior of the SOM imposed by $\rho_n(k)$, let $H = \{h_1, h_2, h_3\}$ be the set of BMUs that will include n on its neighborhood, and thus, affect it. Suppose that during the first epoch and for the first mapped vector v_1, h_1 becomes the BMU. That is, n will be affected by a factor $\rho_n(h_1) = f(1)$ as $\Omega_n(h_1) = 1$. Suppose now that h_2 becomes the BMU for another vector, say v_2. As n is also in the neighborhood of h_2, it will be affected by a factor of $\rho_n(h_2) = f(0.5)$ because $\Omega_n(h_2) = 1/2$, as n has been included in any BMU's neighborhood twice ($T_n = 2$). Now, suppose h_3 becomes the BMU for v_3. It will affect n by $\rho_n(v_3) = f(1/3)$ as $\Omega_n(h_3) = 1/3$. If h_1 happens to be the next BMU, it will affect n with a strength $\rho_n(\Omega_n(h_1)) = f(2/4)$, as n has been affected four times ($T_n = 4$) and h_1 has affected it two times ($T_n^{h_1} = 2$).

Those BMUs that include n in its neighborhood with a high frequency will have a higher effect than the effect of a BMU that includes n as its neighbor less frequently. So, as long as the neighborhood is large enough as to include n, the neurons in H will affect it as a function of $\Omega_n(h_i)$.

Given the case that, after neuron n has been affected for T times a new BMU, say, h_k enters H, it will affect n with a very low intensity ($f(1/T)$), even if n is closer to h_k than to the previous BMUs in H. This modification leads to a momentum effect, in which a neuron tends to be influenced with high intensity by BMUs that have affected it more often than those BMUs that affect it with a low frequency. The more a BMU affects a neuron, the more that neuron will tend to be linked to the BMU: $\rho_n(k)$ establishes the synaptic strength from neuron k to neuron n.

In the SOM, a neuron r (see fig. 1) may be affected by BMUs i and j and its weight vector would be modified with the same magnitude, as long as the neighborhood is large enough as to include it on the influence area of both BMUs even though the distance from i to r is greater than the distance from j to r. This magnitude is independent on the annealing scheme of $\alpha(t)$, as this annealing scheme

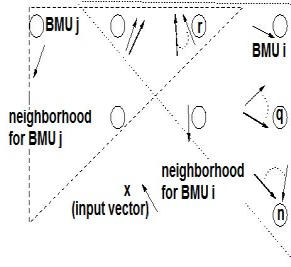is only a function of time and the previous learning factor value, $\alpha(t-1)$ [16].

Figure 1: The influence, shown here as a rotation in the weight vector, of BMU i is higher on neuron q. Thick arrows represent weight vector after modification.

In SOM during the same epoch both r and q are equally affected by i, as they are located at the same distance from it and the learning factor for i is the same. In AFSOM, the influence from i to r is modified with respect to the influence shown in the SOM because j has affected r. Even when the neighborhood function decreases as to exclude r from j's neighborhood, the previous influence is recorded in r's memory ($\rho_r(.)$), so it will not be affected by i with the same intensity as if r had not been affected by j. Also i does not affect r and q with the same magnitude because r has been under the influence of other BMUs (j), which is enough to change the activation frequency by which r is affected by i: no matter if j does not affect r any longer, the activation frequency from i to r has taken that influence into account.

As $\Omega_n(k)$ depends on the number of times k include n on its influence area, $\rho_n(k)$ depends on the neighborhood function and thus it would be decreasing with time, since the neighborhood is also decreasing with respect to time. Therefore the annealing scheme for $\rho_n(k)$ is straightforward as long as it is positive definite.

The form of $\rho_n(k)$ is critical: it should be a positive definite function on $\Omega_n(k)$, but that is not enough. Empirically, we have found that those functions whose image is in the range $[0.5, 1.0]$ not only show lower errors, but also form activity areas which are not symmetrical.

A naive choice for activation frequency function would be $\rho_n(k) = \Omega_n(k)$. This function forms maps whose errors are similar to the maps formed by the SOM, but at the same time, it shares the same properties that other activation frequency functions with lower errors. As the number of epochs go by, the area affected by $\rho_n(k)$ is reduced, as the neighborhood function is decreasing, but the intensity over the neurons being affected by k is higher. However, this intensity is not radial as some of the neurons affected by k could be affected by other neurons.

In the AFSOM every neuron has a temporal memory, $\Omega_n(k)$, which allows it to recall the number of times any other neuron has affected it. This memory could be stated as the synaptic strength between every pair of neurons. If

a given BMU persistently affects a neuron, and no other BMU affects it, then its synaptic strength is higher than the synaptic strength between another neuron and the two or more BMUs that affect it.

Several activation frequency functions were studied in order to identify the general properties that allow: 1) low error measures, 2) low sensitivity to initial conditions and 3) non radial influence areas for BMU.

From all the possible functions for $\rho_n(k)$, only a few of them show error measures lower than the SOM and achieved those low errors through non-radial influence areas. Rule 1, despite of its simplicity, shows error measures similar to the maps trained by SOM with bubble neighborhood, whereas the remaining three rules perform better.

$$(Rule\ 1) \qquad \rho_n(k) = \Omega_n(k) \qquad (3)$$

$$(Rule\ 2) \qquad \rho_n(k) = \frac{1}{1 + e^{-2.5 \times \Omega_n(k)}} \qquad (4)$$

$$(Rule\ 3) \qquad \rho_n(k) = 0.5 \times \Omega_n(k) + 1 \qquad (5)$$

$$(Rule\ 4) \qquad \rho_n(k) = \frac{1}{1 + e^{-0.5 \times \Omega_n(k)}} \qquad (6)$$

The range of rules 2-4 is $[0.5, 1]$ while the range for rule 1 is $[0, 1]$. In several experiments with dozens of different activation frequency functions, those with low errors and non-radial influence were those whose range is $[0.5, 1]$ and are positive definite over $\Omega_n(k)$. This means that the influence from a BMU to neurons on its neighborhood should not be very low in any case, even if it is affecting them for the first time.

In SOM, the neighborhood function defines an influence area of the BMU. It defines a kind of neuronal activity spread effect [17, 27]. The intensity in the influence area is homogeneous, while in AFSOM, those influence areas are neither radial nor symmetrical and, at the same time, the formed maps are as good (and in many cases better) as those formed by SOM with bubble neighborhood, in the sense of low error measures.

If $\rho_n(.) < 1$ then neuron n will not modify its weight as much as stated by $\alpha_n(.)$ and by $|w_n - x|$. The lower $\rho_n(.)$ the lower the modification of n's weight vector in direction of the input vector. That is, n's weight vector will not be attracted to the input vector as much as if this influence were maximal ($\rho_n(.) = 1$). Thus n is more likely to be linked to (affected by) a BMU k if there is a long-time effect (a high value of $\rho_n(k)$), rather than linked to a BMU j that has a recent influence over it (see fig. 2).

In the SOM, the BMU updates its weight vector so it is the closest one to the input vector that stimulated the BMU. Although some other neurons may update its weight vector as much as the BMU, no one can be more similar to the
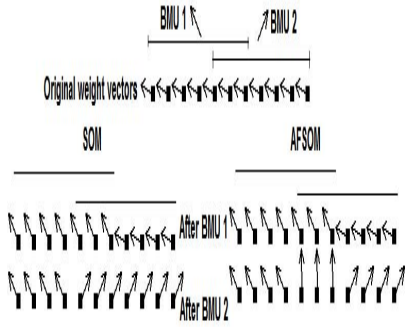
Figure 2: Weight vector modification after two input vectors are mapped to BMUs for SOM and for AFSOM for $\alpha = 1$ (left) and $\alpha = 0.5$ (right). It is observed that for BMU 2 the weight vectors follow a different route in AF-SOM for the three central neurons, as they are under the influence of two BMUs. The neighborhood for each BMU is indicated as a solid line.
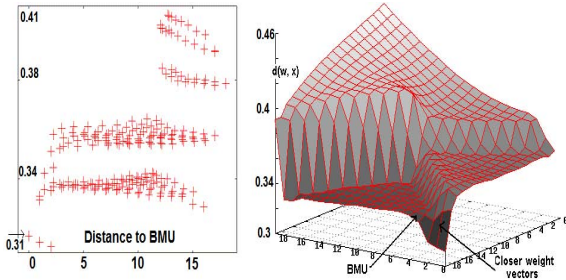

Figure 3: a) Distance from weight vector to input vector (y axis) and distance from the BMU to its neighbors (x axis). The BMU's weight vector (BMU at distance 0) is not the closet to the input vector after applying eq (2). The arrows show the BMU. b) The z axis is the Euclidean distance from each neuron's weight vector to input vector.

input vector than BMU's because of eq. (1). In contrast, in AFSOM, the BMU's weight vector may not necessarily become the most similar one to the input vector (see fig 3). That is, other neurons may update their weight vector and become more active that the BMU, in the sense of similitude to the input vector.

Let k be the BMU for a certain input vector m and neuron i to be in its neighborhood. It is possible that the weight vector for i gets closer to m than k's weight vector. To see this, lets consider the adaptation process that takes place after input m is mapped to k: $w_k = w_k + \rho_k(k)(m - w_k)$. Also, neuron i is updated: $w_i = w_i + \rho_i(k)(m - w_k)$ (the time t has been eliminated for clarity and α and h are not indicated as both have the same value for i and k in the bubble neighborhood scheme). If it happens that:

- $\rho_k(k) + \delta < \rho_i(k)$ as, for example, if neuron i has not been affected by any other neuron different from k whereas k has been influenced by other neurons, and
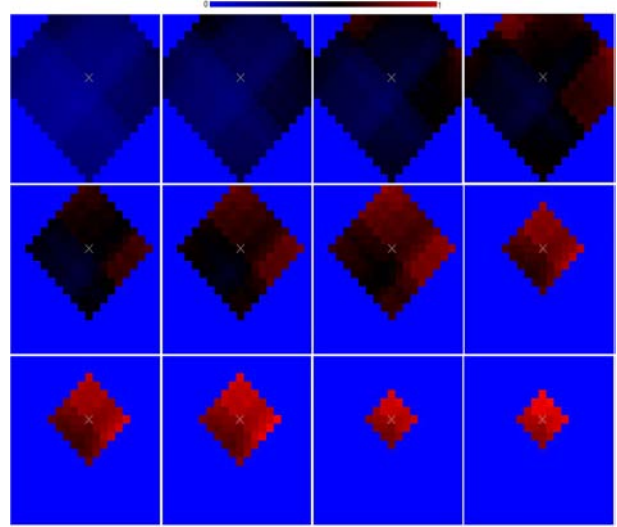

Figure 4: Influence areas for a BMU (white x) for rule 1 for four consecutive epochs. The intensity of the influence from the BMU to its neighbors is not radial or symmetrical. As the number of epochs and mapped vectors increases, the intensity of the influence increases. In a-d (starting at top left), four different input vectors were mapped to the BMU in the same epoch. It is observed a bias in the intensity. The color scale is proportional to $\rho_n(k)$.

- $|w_k - m| < |w_i - m| + \delta$ which means that weight vector for neuron i is not "very far away" from m input vector

then weight vector i gets closer to m than weight vector k as $\rho_k(k)|w_k - m| < \rho_i(k)|w_i - m|$.

The activity patterns formed by $\rho_n(k)$ show interesting properties. For example, neurons that are very close to BMU k are not affected as much as some neurons that are located far away from k, as it may be observed in figs. 4c)-e), where close neurons are blue-colored, while farther neurons are red-colored, which means that $\rho_n(k)$ is higher for the later neurons. This is explained by the fact that the blue-colored neurons are being affected with a high frequency by other BMU's different from k, while the red-colored are being affected with a low frequency (if any) by other BMUs.

The training of the map in the AFSOM is affected by $\rho_n(k)$. Fig 5 shows the weight folding for SOM and for the proposed modifications for an artificial data set. It is observed that the AFSOM leads to a faster folding than SOM.

3 Results

We evaluated the AFSOM for six data sets: the two-dimensional spiral (350 vectors), the iris data set, the ionosphere data set, the codon usage data set (64-dimensional and 400 input vectors), the waveform data set and the Mexi-
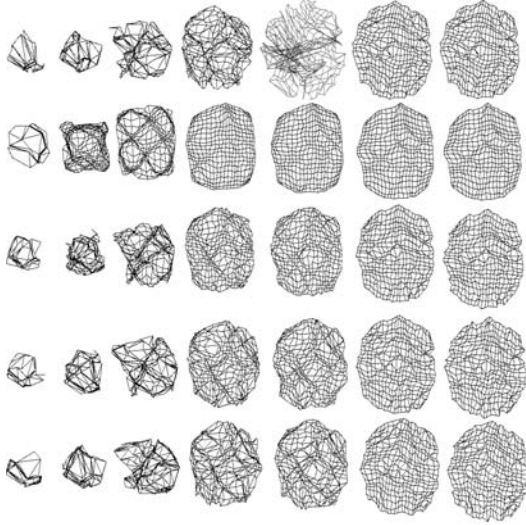
Figure 5: Weight folding for epochs $t = 1, 3, 6, 10, 12, 20$ and 30 for SOM (first row) and for the AFSOM (equations (3)-(6), rows 2 - 4) obtained in a 20×20 lattice. The data set is an unitary circumference with 1000 input vectors.

can elections (ME) data set, (six-dimensional and 300 input vectors).

In order to verify low-error maps formed by eq. (2) two error measures were quantified and compared to the error measures present in the maps formed by eq (1). Although there are several error measures for the maps obtained by the SOM algorithm, [10, 15, 14], the topographic error (TE) as well as the error quantization (EQ) were the used error measures for the obtained maps, as they are good measures of the quality of topographic mapping and vector quantization. In order to test sensitivity and self-organization, several thousands of experiments were made for three different lattices size, $N \times N$ ($N = 10, 20$ and 25), as well as for the initial learning parameter $0 < \alpha(0) \leq 1$ and for the initial neighborhood size $1 < h_n(g, 0) \leq N$. Practically, for each

- learning set (spiral, iris, ionosphere, ME data, codon usage and waveform),
- number of epochs (between 1 and 30),
- activation frequency function (equation (1) was also included for comparison),

the initial learning parameter $\alpha(0)$ was chosen randomly from $(0, 1]$ as well as the initial neighborhood width was chosen from $[1, N]$. The final learning parameter was 0.00001. The final neighborhood values were set to 0 and decreased exponentially.

Figures 6 - 7 show TE and EQ as a function of the number of epochs and initial neighborhood width for the SOM and for AFSOM for the six data sets for the 20x20 lattice. Rules 1 and 3 show similar error measures for the six data sets. Both of them show low TE values if both, the number of epochs and initial neighborhood width are low. SOM
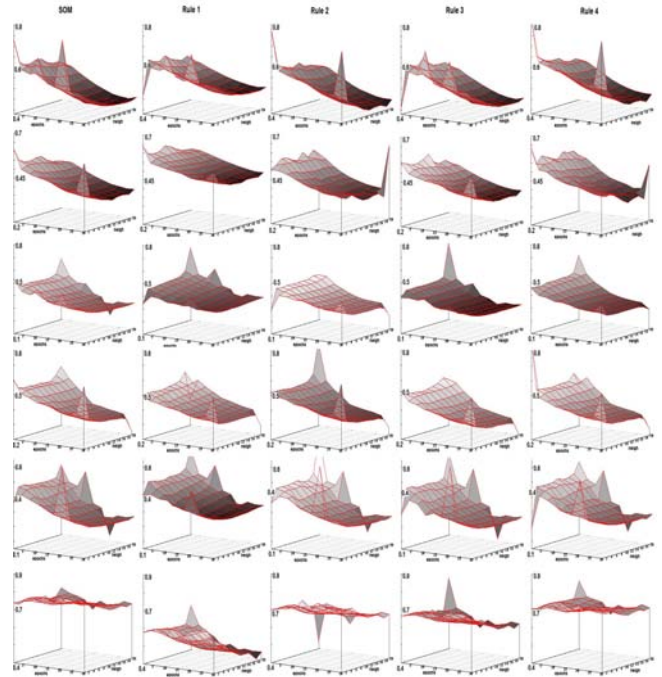

Figure 6: TE as a function of number of epochs and initial neighborhood width for the six data sets (from top to bottom spiral, iris, Mexican elections, ionosphere, codon usage and waveform) for SOM (first column) and AFSOM (rules 1 - 4, columns 2 - 5).

shows high TE values in the former situation. Rule 1 shows very low TE for the waveform data set if compared to any other rule, including SOM.

Rules 2 and 4 show also similar TE curves. For the iris data set, both rules show high TE values in an unexpected case, when the number of epochs and the initial neighborhood width are maximal, which contrasts with the behavior shown by SOM and rules 1 and 3. For EQ, rules 1 and 3 show lower values than SOM, with exception of the ME data set, for which rule 3 shows high values. EQ seems harder to be decreased in the AFSOM for all data sets.

Tables (1)-(3) show a summary of the TE and EQ for the analyzed data sets. It is shown that the topographic error is, in general, lower in the maps obtained by eq (2) than in those maps obtained by eq (1). Consistently, rule 3 forms maps with lower TE and EQ values than those formed by the remaining rules as well as those formed by SOM.

4 Conclusions

A parameter that accounts for the activation frequency between BMUs and its neighbors is incorporated to the SOM. In it, each neuron has a memory of the synaptic strength it has with each other neuron in the network. The activation frequency resembles Hebbian learning, as it is a function of the frequency by which neurons respond for the same
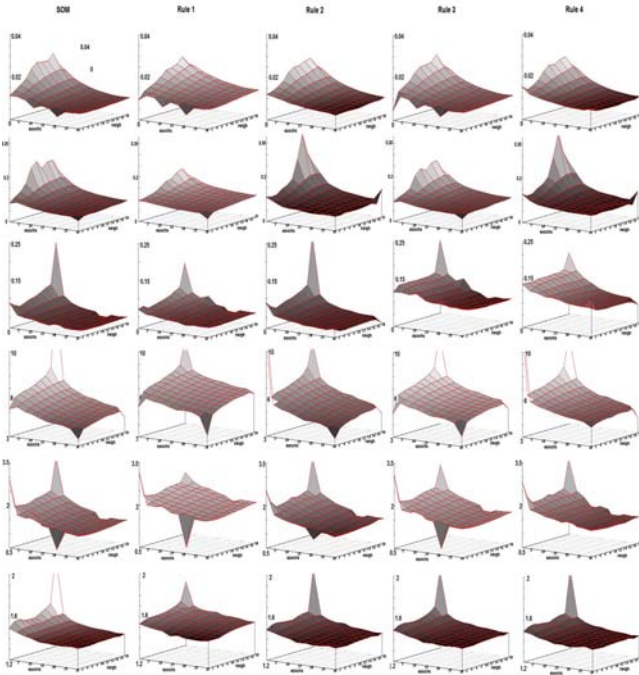
Figure 7: EQ as a function of number of epochs and initial neighborhood width for the six data sets for SOM (first column) and AFSOM (rules 1 - 4, columns 2 - 5).

stimulus, achieved by a differentiated influence of BMU to its neighbors.

The weight folding in the proposed model is subject to the differentiated influence each BMU k has in its neighbors. The weight vector for those neurons with which k has a strong synaptic strength will be strongly attracted to the input vector for which k became BMU. This behavior allows a faster weight folding and also allows the weight vector from other neurons to get closer to the input stimulus than the BMU's weight vector.

The fact that the BMU's weight vector may not be the closest to the input stimulus is also remarkable. As it may happen that a BMU affects with high intensity neurons that are far away and affect with lower intensity neurons closer to it, it may explain the brain cortex images obtained by neuroimaging techniques in an alternative way. Research has been done to verify the compartmental structure of the brain and it has been proposed that connectivity in the brain may have a strong genetic nature [13, 28].

However, with the proposed model, it may be possible to explain those images in a different way. Here, the neurons that are active for a particular stimulus may not be the most excited ones, if excitation is equivalent to distance from weight vector to input vector after adaptation. So, if a given neuron responds to a stimulus, it may not even be very excited, because of the frequency activation function, but may be able to stimulate neurons located far away, and in a non-radial pattern. Following this idea, the neurons identified as stimulated may not necesarily be the most excited ones for that stimulus.

Although the error measures are in general lower in AF-SOM than in SOM, this is not its main feature. The fact that it achieves self-organization by a non-radial influence of neurons is more important, because it resembles more accurately the brain cortex dynamics by generating non-symmetrical activity patterns and also because it incorporates memory between every pair of neurons.

Table 1: Mean (μ) and standard deviation (σ) for 20 000 experiments with random initial neighborhood width ($1 < h < 20$), random inital α and random number of epochs ($0 < r < 31$), for the spiral and iris data sets, for 20x20 lattice. $h(r) = 0$ and $\alpha(r) = 0.00001$.

Spiral				
R.	μTE	σTE	μEQ	σEQ
SOM	0.50	0.001	0.010	0.01
1	0.50	0.000	0.011	0.01
2	0.49	0.001	0.009	0.01
3	0.49	0.002	0.010	0.02
4	0.49	0.001	0.010	0.01
Iris				
R.	μTE	σTE	μEQ	σEQ
SOM	0.41	0.001	0.092	0.000
1	0.42	0.001	0.091	0.000
2	0.41	0.001	0.092	0.001
3	0.40	0.001	0.091	0.001
4	0.40	0.001	0.092	0.000

Table 2: μ and σ for 20 000 experiments with random initial neighborhood width and random inital α for the Mexican elections and ionosphere data sets, for 20x20 lattice.

Elections				
R.	μTE	σTE	μEQ	σEQ
SOM	0.36	0.001	0.049	0.002
1	0.36	0.008	0.051	0.001
2	0.35	0.009	0.049	0.003
3	0.34	0.006	0.049	0.002
4	0.36	0.001	0.048	0.003
Ionosphere				
R.	μTE	σTE	μEQ	σEQ
SOM	0.42	0.003	5.347	0.007
1	0.41	0.001	5.785	0.005
2	0.41	0.002	5.329	0.007
3	0.38	0.001	5.348	0.008
4	0.41	0.001	5.667	0.010

Table 3: μ and σ for 20 000 experiments with random initial neighborhood width and random α for the codon and waveform data sets, for 20x20 lattice.

	codon			
R.	μTE	σTE	μEQ	σEQ
SOM	0.35	0.003	2.307	0.024
1	0.37	0.005	2.408	0.032
2	0.38	0.004	2.394	0.081
3	0.33	0.001	2.306	0.028
4	0.37	0.007	2.537	0.091
	waveform			
R.	μTE	σTE	μEQ	σEQ
SOM	0.73	0.001	1.441	0.002
1	0.52	0.004	1.501	0.000
2	0.63	0.001	1.447	0.001
3	0.68	0.001	1.432	0.001
4	0.70	0.001	1.442	0.002

References

[1] Amari, S. Topographic organization of nerve fields. Bulletin of math. Biol. 42 (1980) 339-364.

[2] Bednar, J., Kelkar, A, Miikkulainen, R. Scaling self-organizing maps to model large cortical networks. Neuroinformatics. (2004).

[3] Cabeza, R., Kingston, A. Handbook of functional neuroimaging of cognition. MIT Press. (2006).

[4] Chappell, G., Taylor, J. The temporal Kohonen map. Neural Networks 6 (1993) 441-445.

[5] Cottrell, M. Fort, J.C., Pagés, G. Theoretical aspects of the SOM algorithm. Neurocomputing 21 (1998) 119-138.

[6] Erwin, Obermayer, K. Schulten, K. Self-organizing maps: Ordering, convergence properties and energy functions. Biol. Cyb. 67 (1992) 47-55.

[7] Flanagan, J. Sufficiente conditions for self-organization in the SOM with a decreasing neighborhood function of any width. Conf. of Art. Neural Networks. Conf. pub. No. 470 (1999).

[8] Gerstner, W., Kistler, W. Mathematical formulations of Hebbian learning Biol. cyb. 87 (2002) 404-415.

[9] Goodhill, G., Cimponeriu, A. Analysis of the elastic net model applied to the formation of ocular dominance and orientation columns. Network: comput. neural systems 11 (2000). 153-168.

[10] Goodhill, G., Finch, S., Sejnowski, T. Quantifying neighborhood preservation in topographic maps. 3rd. joint Symp. in neural computation (1996) 61-82.

[11] Hammer, B., Micheli, A., Sperduti, A., Strickert, M. Recursive self-organizing network models. Neural Networks 17 (2004) 1061-1085.

[12] Kamimura, R. Cooperative information control for self-organizing maps Neurocomputing 62 (2004) 225-265.

[13] Kandel, E., Schwartz, J., Jessell, T. Principles of neural science. McGraw-Hill. (2000).

[14] Bauer, H., Herrmann, M., Villmann, T. Neural maps and topographic vector quantization. Neural networks 12 (1999) 659-676.

[15] Kiviluoto, K. Topology preservation in Self-Organizing maps. Proc. ICNN96, IEEE Int. Conf. on Neural Networks. (1996).

[16] Kohonen, T. Self-Organizing maps. 3rd. ed. Springer-Verlag. (2000).

[17] Kohonen, T. Self-organizing neural projections. Neural Networks 19 (2006). 723-733.

[18] Koulakov, A., Chklovsky, D. Neuron 29 (2001) 519-527.

[19] Lee, J., Verleysen, M. Self-organizing maps with recursive neighborhood adaption. Neural Networks 15 (2002) 993-1003.

[20] Liljeholm, M., Lin, A., Ozdzynski, P., Beatty, J. Neurocomputing. 44-46 (2002) 515-520.

[21] Malsburg, C. Self-organization of orientation sensitive cells in the striate cortex Kybernetic 14 (1973) 84-100.

[22] Mayer, N., Herrmann, J., Geisel, T. Neurocomputing 32-33 (2000) 447-452.

[23] Miikkulainen, R. Bednar, J. Choe, Y. Sirosh, J. Computational maps in the visual cortex. New York. Springer (2005).

[24] Neme, A., Miramontes, P. ICANN (1) (2005) 427-432.

[25] Neme, A., Miramontes, P. ICANN (1) (2006) 455-463.

[26] Principe, J., Euliano, N., Garani, S. Principles and networks for self-organization in space-time Neural Networks 15 (2002) 1069-1083.

[27] Euliano, N., Principe, J. A spatio-temporal memory based on SOMs with activity diffusion. WSOM. 253-266. (1999).

[28] Rakic, S., Zecevic, N. Emerging complexity of layer I in human cerebral cortex. Cerebral Cortex 13 (2003) 1072-1083.

[29] Ritter, H. Self-Organizing Maps on non-euclidean Spaces Kohonen Maps, 97-108, Eds.: E. Oja and S. Kaski, (1999).

[30] Sakata, S., Yamamori, T. Topological relationships between brain and social networks Neural networks 20 (2007) 12-21.

[31] Sporns, O., Chialvo, D., Kaiser, M., Hildetag, C. Organization, development and function of complex brain networks. Trends in cog. sciences 8 No. 9 (2002) 418-425.