Applied Mathematical Sciences, Vol. 8, 2014, no. 100, 4953 - 4969
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2014.310585

Self-Organizing Maps and Entropy Applied to Data Analysis of Functional Magnetic Resonance Images

Anderson D. S. Campelo, Valcir J. C. Farias, Héliton R. Tavares and Marcus P. da C. da Rocha

PPGME-ICEN-Federal University of Pará, Brazil

Abstract

Kohonen self-organizing maps (SOM) and Shannon entropy were applied together for the analysis of data from functional magnetic resonance imaging (fMRI). To increase the efficiency of SOM in the search for patterns of activation in fMRI data, first, we applied the Shannon entropy in order to eliminate signals possibly related to noise sources. The procedure with these techniques was applied to simulated data and on real hearing experiment, the results showed that the application of entropy and SOM is a good tool to the identification of areas of activity.

Keywords: Shannon entropy, functional magnetic resonance imaging, self-organizing maps

1. INTRODUCTION

Functional magnetic resonance imaging (FMRI) is a non-invasive tool widely used for studying the human brain in action. The FMRI has been applied to cognitive studies and also in a clinical setting to monitor tumour growth, pre-surgical mapping, mental chronometry studies and to diagnose epilepsy, Alzheimer's disease, etc [18]. The FMRI measurements are based on blood-oxygen-level-dependent correlations (BOLD) [23,12], with hemoglobin being used as endogenous contrast agent, due to the magnetic properties of oxy–hemoglobin (diamagnetic) and deoxy–hemoglobin (paramagnetic)[13].

The BOLD signal can be obtained using two experimental paradigms. The first used a blocked design, with the subject being exposed to alternating periods of

stimulation and rest. The event-related paradigm, on the other hand, required that the subject performs a simple task, intercalated by long resting intervals.

A fMRI dataset consists of images in 3D space ($x \times y \times z$), with each image point, named a voxel, changing a long time (t). Most fMRI analysis try to identify how signal related to voxels in a region of interest (ROI) vary in time and to find out whether these variations are somehow correlated with the stimulus. This analysis, however, is a computational challenge due to the low signal to noise ratio in the BOLD response and the usually large amount of data that needs to be processed. Many analytical methods have been developed to deal with this complexity, some of them were created earlier to analyze positron emission topography-generated signals (PET).

Most methods available in the literature use statistical techniques to identify active regions, including Student´s t test [20], crossed correlation [24], and the general linear model (GLM)[7]. These methods are based on the standard hemodynamic function, which models the BOLD response in the brain. Other popular methods are independent component analysis (ICA)[16,1] and principal component analysis (PCA)[8] and there are works apply entropic method over FMRI data with event-related paradigm [3].

Grouping techniques have also been used successfully, including K-means [2], fuzzy cluster [17] and hierarchical clustering [14]. Clustering techniques are based on the similarity observed in voxel's time series.

Within the same class of methods, one can still highlight the self-organizing maps (SOM), a type of artificial neural network (ANN) for visualization and clustering of high dimensional data that preserves a spatial organization using a competitive learning [26]. The SOM has been applied to fMRI data analysis by Fischer and Henning [6]. Ngan et al. [21] used the method together with the cross-correlation analysis to identify regions associated with the visual and motor stimulus. Peltier, et al. [22] applied the SOM to identify functional connectivities of low-frequency at fMRI data. Chen et al. [5] and Liao et al. [28] used SOM associated with a hierarchical clustering to detect different experimental designs on the same map.

To increase the efficiency of the method used for fMRI Data analysis is recommended to reduce the number of voxels due to the fact that the proportion of active voxels on the brain is relatively small compared to the number of voxels that are not active [19].

In this paper a study on self-organizing maps and hierarchical clustering techniques as tools for analysis of fMRI data will be made. The SOM will be associated with the Shannon entropy in order to make it more effective in detecting patterns related to brain activity. To measure the efficiency of the SOM, after

applying this in conjunction with the entropy is calculated the quantization error, a measure that evaluates the quality of the highest brightness intensity of input data. Here, some voxels will be dropped from the application of Shannon entropy in the fMRI data.

2. MATERIAL AND METHODS

2.1 Analysis of Entropy

Some authors have recommended reducing the number of voxels in the application of clustering algorithms, due to the fact that the proportion of active voxels in the brain is relatively small [19]. So to increase the efficiency of data analysis, only signals corresponding to the brain were processed. Moreover, it was applied to each voxel an analysis of entropy using Shannon entropy, to eliminate some positions on the vector of input data. The voxels are eliminated with a level of entropy below a certain threshold determined empirically after several tests.

The calculation of the entropy of the signal was based on work by De Araujo et al. [3], where *S(t)* denotes the behavior of the time series of a voxel (hemodynamic response-HRF). The signal is divided into two levels of intensity $(I_l,\ l = 1,2)$. This division is performed by calculating the maximum and minimum values of the time series.

Mathematically, consider D as a discrete set of amplitudes sampled in time:

$$D = \{s(t_k), k = 1, \dots, K\} \tag{1}$$

Taking,

$$s_0 = \min_k\{s(t_k), k = 1, \dots, K\}, \tag{2}$$

$$s_L = \max_k\{s(t_k), k = 1, \dots, K\}. \tag{3}$$

Thus, the amplitude levels I_1 and I_2 are defined as

$$I_1 = [s_0, s_M) \tag{4}$$

$$I_2 = [s_M, s_L] \tag{5}$$

where s_M is defined as the arithmetic mean between s_0 and s_L

Denote by $P(I_l)$ the probability of signal $s(t_k)$ to belong to the interval I, which can be calculated by

$$P(I_l) = \frac{1}{K}\sum_{k=1}^{K}\delta_{\mathrm{IN}(k)}, \tag{6}$$

where $N(k)$ is the level at which the amplitude $s(t_k)$ belongs and $\delta_{lN(k)}$ is given by:

$$\delta_{lN(k)} = \begin{cases} 1, & se\ I_l = N(k), \\ 0, & se\ I_l \neq N(k). \end{cases} \tag{7}$$

The Shannon entropy associated with this set of probabilities can now be calculated according to the equation:

$$H = -\sum_{l=1}^{2} P(I_l) \log_2 P(I_l). \tag{8}$$

The Figure 1 shows the Shannon entropy for two levels of intensity, ie, in the case of two possibilities with probabilities p and $(1-p)$. Looking at Figure 1, it can be seen that the Shannon entropy reaches a maximum when the possibilities are evenly distributed and it reaches a minimum when has the distribution concentrated in any one of the possibilities.

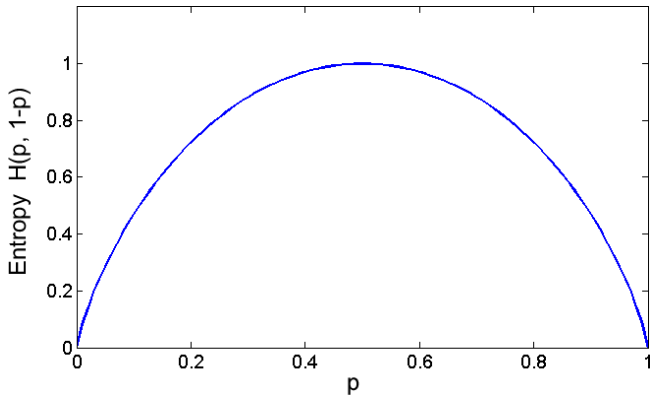


Figure 1. The Shannon entropy in the case of two possibilities with probabilities p and $(1-p)$.

Consider two time series originated from an fMRI experiment, a signal representing a non-active voxel (Figure 2 a) and another representing the sign of an active voxel. In Figure 2a notes that for a signal corresponding to a non-active voxel, the presence of spurious peaks makes the probability distribution narrow, therefore, a low entropy value. In Figure 2b, the time series of an active voxel, the high value of entropy is associated with a wide probability distribution.

2.2 Self-Organizing Maps

FMRI data was analyzed with Kohonen´s SOM [26] using an implementation available in the literature [6, 21, 22]. Kohonen´s SOM is an artificial neural network where neurons are disposed as a uni- or bi-dimensional grid layout. In a bi-dimensional layout, the geometry is free and can be rectangular, hexagonal, triangular etc. In a SOM, each neuron in a grid is represented by a probability distribution function of the input data.

The SOM algorithm responsible for map formation begins initializing the grid neurons weights with random values, which can be obtained from the input data. In the present work, it was used a bidimensional grid of dimension 10×10 [22]. Each neuron in the grid is connected to every element of the input dataset, i.e., the dimension of weights $\mathbf{m}_i$ (i=100) is the same as the input dataset:

$$\mathbf{m}_i = [m_{i1}, m_{i2}, \ldots, m_{in}]^T \in \Re \tag{9}$$

where n indicates the total amount of points available in the time series generated by the fMRI experiment.

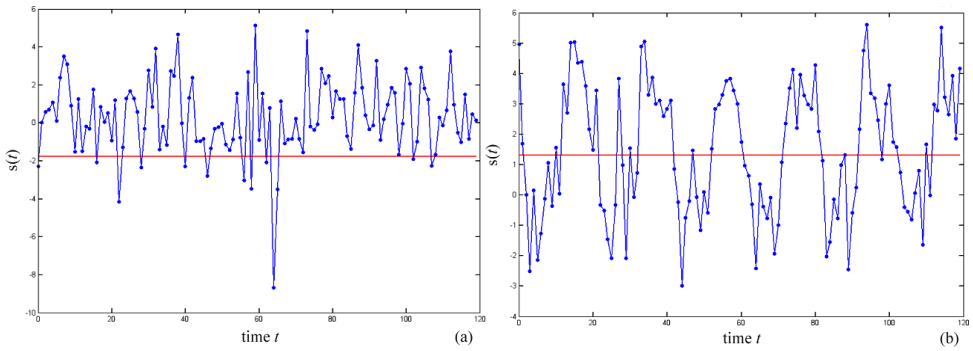


Figure 2. (a) Signal with low entropy, the points are concentrated in the up interval, with the following probability distributions of each interval, $P(I_1) = 0.1083$ e $P(I_2) = 0.8917$ and entropy $H = 0.4949$; (b) Signal with high entropy, with probability distribution for the intervals $P(I_1) = 0.4667$ e $P(I_2) = 0.5333$ and entropy $H = 0.9968$.

After each iteration t of the ANN, we selected randomly a vector from the input dataset, given by:

$$\mathbf{x}_i = [x_1, x_2, \ldots, x_n]^T \in \Re \tag{10}$$

which indicates the time series of a given voxel from the fMRI dataset.

Then x is compared to all the weights of the grid, with frequent use of the minimum Euclidean distance as similarity criterion for choosing a winner neuron [6, 22]. However, the correlation as a similarity criterion reveals itself better than the conventional Euclidean distance [27], the choice of the winning neuron c given by:

$$\mathbf{m}_c = \arg\max_i\{\mathrm{corr}(\mathbf{x}(t), \mathbf{m}_i(t))\} \tag{11}$$

with $i = 1, \ldots, M$ where M the total number of neurons in the grid, $\mathbf{m}_c(t)$ represents the time series of the winner c and $\mathrm{corr}(\mathbf{x}(t), \mathbf{m}_i(t))$ is the correlation coefficient between $\mathbf{x}(t)$ and $\mathbf{m}_i(t)$.

The updating of the weight vector $\mathbf{m}(t+1)$ in time $t+1$, with t= $0, 1, 2, \ldots$ is defined by:

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + h_{ci}(t)[\mathbf{x}(t) - \mathbf{m}_i(t)] \tag{12}$$

which is applied to every neuron on the grid that is within the topological neighborhood-kernel h_{ci} from the winner neuron c.

The equation (12) has the goal of approximating the weight vector $\mathbf{m}_i$ of neuron i towards the input vector, following the degree of interactions h_{ci}. This approach transforms the grid, after training, in a topologically organized characteristic map, in the sense that adjacent neurons tend to have similar weights.

A function frequently used to represent the topological neighborhood-kernel h_{ci} is the Gaussian function, which is defined by:

$$h_{ci}(t) = \alpha(t) \exp\left\{-\frac{\|\mathbf{r}_c - \mathbf{r}_i\|}{2\sigma^2(t)}\right\}, \tag{13}$$

where $\alpha(t)$ is the learning rate, which has to gradually decrease along time to avoid that new data gathered after a long training session could compromise the knowledge already sedimented in the ANN; $\mathbf{r}_c$ and $\mathbf{r}_i$ determine the discreet position of neurons c and i in the grid and $\sigma(t)$ defines the topological neighborhood radius, i.e., defines the full-width at half-maximum (FWHM) of the Gaussian kernel. The parameters $\alpha(t)$ and $\sigma(t)$ gradually decrease by t/τ (τ is a time constant) after each iteration t, following an exponential decay.

After the learning process of the SOM, the input data with similar patterns appear

in groups in neighboring neurons on the map. However, often the amount of groups is unknown or of complex discrimination, making it difficult to distinguish such groups. A proposal for detection and automatic segmentation of the groups in the SOM map is through clustering techniques. The purpose of the clustering techniques is to form similar groups, ie groups that have a high degree of correspondence or similarity.
Several methods for grouping of neurons in the SOM have been proposed, in this paper will apply the hierarchical clustering (HC), for more details see [27].

2.3 Evaluating the SOM Quality

There are several mechanisms that can evaluate the quality of the generated map obtained after the learning process. In the present work it was used the quantization error:

$$E_q = \frac{1}{N}\sum \|\mathbf{x} - \mathbf{m}_c\|^2. \tag{14}$$

The quantization error is defined as the mean error corresponding to the difference between each characteristic vector $\mathbf{x}$ and the winner neuron $\mathbf{m}_c$, where N is the total number of patterns.

2.4 Simulated Data

It is essential for fMRI Data analysis the knowledge of hemodynamic response function, some the better known functions include the poisson function, the Gaussian function, the inverse logit function, and the linear combination of several functions[4], Although currently the model hemodynamic of Friston [10] be fairly applied, this article will use the linear combination of several functions (Glover), since the interest here is in qualitative data analysis to detect activation and not in the quantitative understanding of physiological factors such as change in flow, oxygen extraction, blood volumes, etc [11, 27, 15].

It was simulated the fMRI synthetic experiment (128×128) with 110 slices, convoluting a block-like stimulus function with the canonical hemodynamic response function generated as a sum of two distribution functions [9]:

$$h(t) = \left(\frac{t}{d_1}\right)^{a_1} \exp\left(\frac{-(t-d_1)}{b_1}\right) - c\left(\frac{t}{d_2}\right)^{a_2} \exp\left(\frac{-(t-d_2)}{b_2}\right) \tag{16}$$

where $d_j = a_j b_j$ is the time until the peak of each gamma function, with $a_1 = 6$, $a_2 = 12$, $b_1 = b_2 = 0.9$, e $c = 0.35$, having been determined experimentally by Glover [4] with auditory stimulation.

The Figure 3a shows two active areas with a total of 524 voxels (A=276 and B=248), while 5176 voxels corresponded to the remaining grey matter. The other 10684 voxels corresponded to the background and are not time modulated. One can observe that the signal of the area A is delayed relative to that of area B (Figure 3b). In time series of the two areas added uniform Gaussian noise to reach a signal noise ratio (SNR) of -5 dB, which was calculated with the following expression:

$$\mathrm{SNR} = 10 \log \left(\frac{\sigma_S^2}{\sigma_R^2} \right) \tag{17}$$

where σ_S^2 and σ_R^2 are signal and noise variances, respectively.

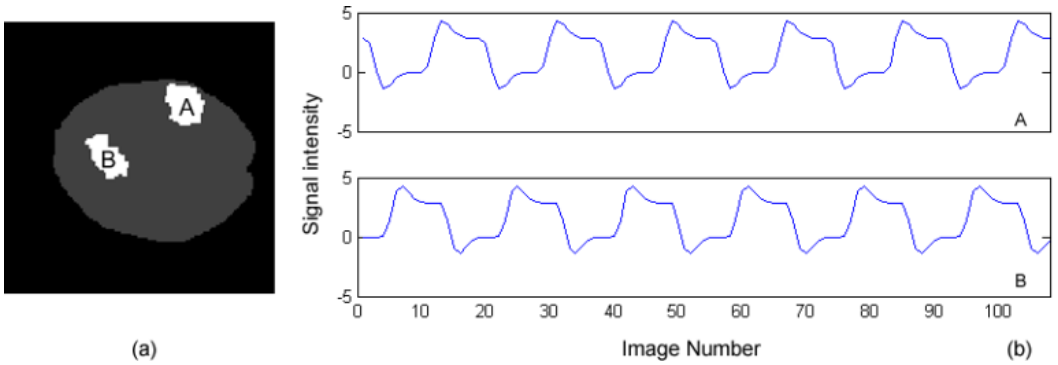


Figure 3. (a) diagram showing the synthetic data. The sub-regions A and B with two different HRFs. (b) Temporal series modeled in sub-regions A and B without adding noise.

2.5 Real Data

The fMRI experiment used a 1.5 T Siemens scanner (Magnetom Vision, Erlangen, Germany), with the following parameters for EPI (echo-planar imaging) sequences: TE = 60 ms, TR = 4.6 s, FA = 90º, FOV = 220 mm, and slice thickness of 6.25 mm. 64 cerebral volumes with 16 slices each were acquired with a matrix dimension of 128 x 128.

During the experimental procedure the subject received auditory stimulation in a blocked design, with 5 stimulation blocks (27.5s each) intercalated with 6 resting blocks (27.5 s each). During the task, the subject listened passively to a complex story with a standard narrative structure. After, the test the subject had to inform to the experimenter its comprehension of the story content.

Acquired images were preprocessed with the software SPM8 (Statistical Parametric Mapping) in order to increase the signal-to-noise ratio (SNR) and to eliminate

incident noise associated with the hardware, involuntary movements of the head, cardiac and respiratory rhythms, etc.

2.6 Procedures

After the preparatory phase of the images (real data), which was performed with the applying of the software SPM8, It were calculated the values of entropy of all voxels of FMRI data, Through the procedure described above. Those voxels that presented value of entropy lower than the chosen threshold were eliminated from the set of input data of SOM.

It was used information obtained in previous studies [22, 28] in both simulated and real data to set the parameters of the SOM. Thus, it was used a two-dimensional rectangular grid 10×10. The learning rate was initialized with the value $\alpha(0) = 0.1$, and the effective width parameter $\sigma(0) = 7$. Although Peltier et al. [22] suggest that 100 iterations is enough, the number of iterations of the SOM was adjusted accord to stabilization of the quantization error (Eq. 14). Alternatively the minimum distance Euclidiana it was applied to maximum correlation between the input data and the neurons of the grid (Eq. 11) for the selection of the winning neuron . After training the SOM, we applied the HC, where only neighboring neurons were clustered on the map.

With the defination of the clusters which show evidence of the BOLD response, a correlation between the mean of the weights of the neurons belonging to the grouping (with evidence of the BOLD response) with all voxels of the images was done. Those voxels that showed a correlation coefficient (CC) with average weights greater than a threshold chosen empirically, are highlighted in a mask with the purpose of generating the map of activations.

3 RESULTS

3.1 Simulated data

The first step of the process was to calculate the Shannon entropy for each voxel of the simulated data. In Figure 4a is possible to observe a high value of entropy for the subregions A and B. The voxels with entropy lower than 0.85 have been eliminated. Figure 4b shows the voxels that were discarded after entropic analysis.

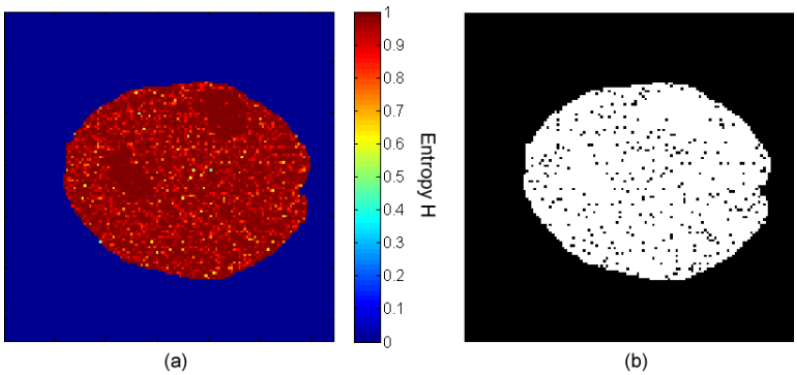


Figure 4. (a) Entropic map of artificial data for the first data set with SNR 5dB. (b) After the entropic analysis of simulated data, the black dots in the image were deleted of the subsequent stage, which corresponds to 6% of the voxels belonging to gray matter.

Later, it was started the training stage of SOM, it has been sufficient a total of 250 iterations, this value was established starting from stabilization of the quantization error. The next step was to detect the neurons that exhibit patterns of BOLD signal. Figure (5a) shows the group of neurons that represent the patterns of activation and were segmented after the application of the HC algorithm in highlighting. It is possible to observe two groups, with color yellow (top left) which corresponds to HRF modeled in the subregion B, while in color green (bottom right) we have the group that represents the modeled HRF in subregion A. The Figure 5b shows the evolution of the quantization error calculated in every 10 iterations, for the two cases with and without the application of the entropy. It is observed that with the entropy the error of quantization reaches the stabilization faster than without the application of the entropy.

The Figures 6a and 6b show the activation map from two groups of neurons highlighted in Figure 5a. The CC between the average of the weights and the signal generated was greater than 0.8 (This value was obtained after several tests on the data)

3.2 Real Data

The Figure 7a shows the entropy analysis for the voxels of the eighth slice of the resonance of the brain whose value of entropy ranged from 0.1161 to 1. It is observed that only with the application of this method it is not possible detect the brain regions responsible for hearing (active regions). Of the total number of voxels inside the

brain, 8% were eliminated from further training of the SOM (Figure 7b). These voxels showed value of entropy $H \leq 0.8$.

The Figures 7c, 7d, 7e show the synaptic weights of the SOM and the quantization error without e with the entropy after 250 iterations, respectively. Analyzing the graph of error is possible to verify that the magnitude of the quantization error for the situation in which the entropy has been applied is smaller, and has begun to stabilize slightly (from iteration 150) earlier than the case in which is not used the entropy(from iteration 200).

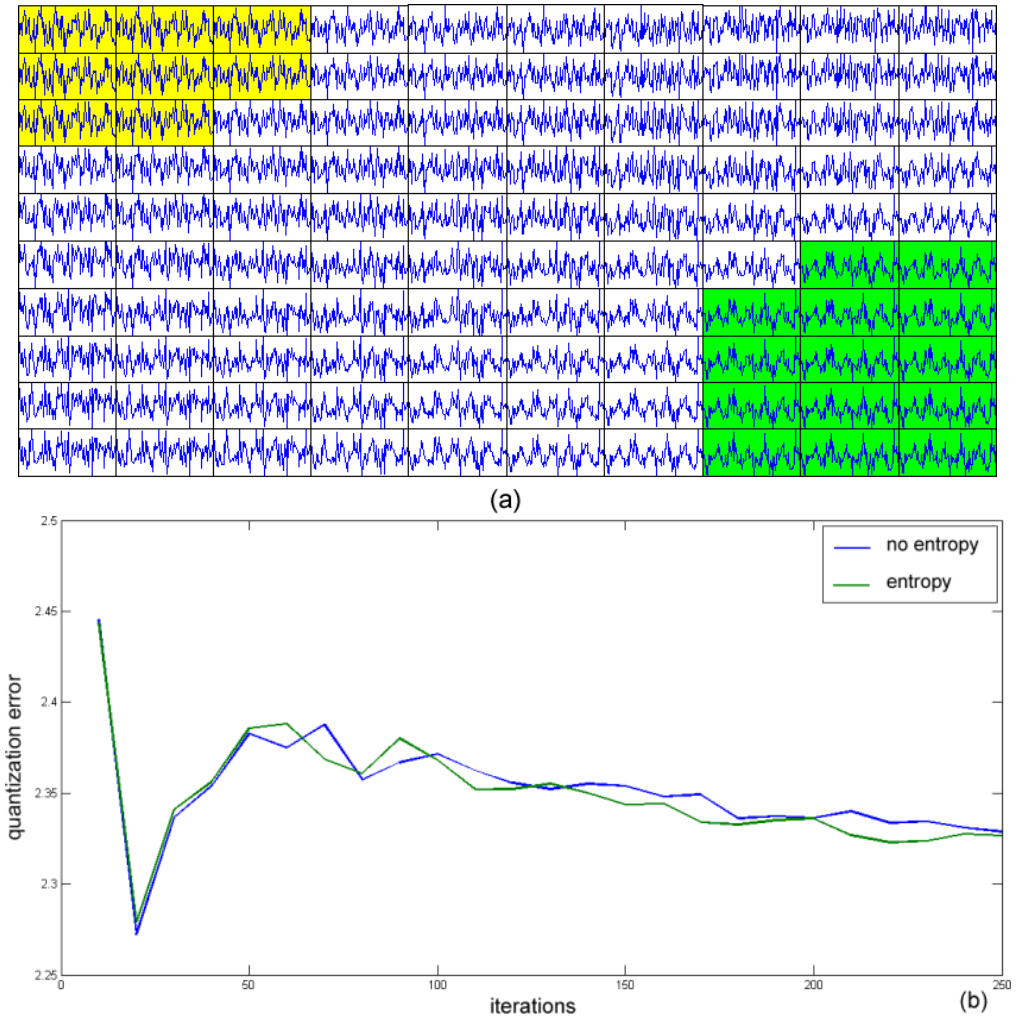
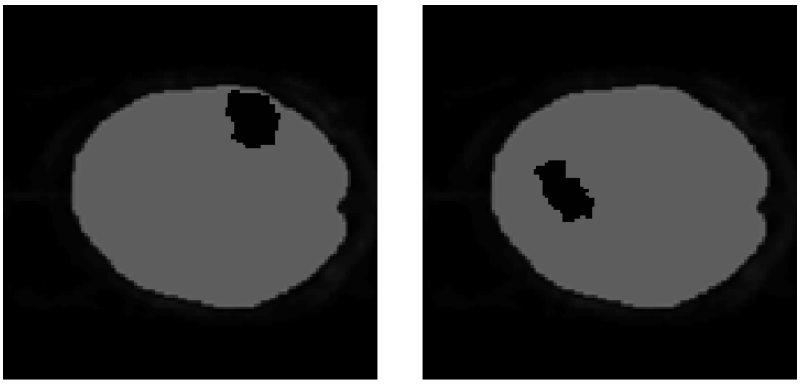
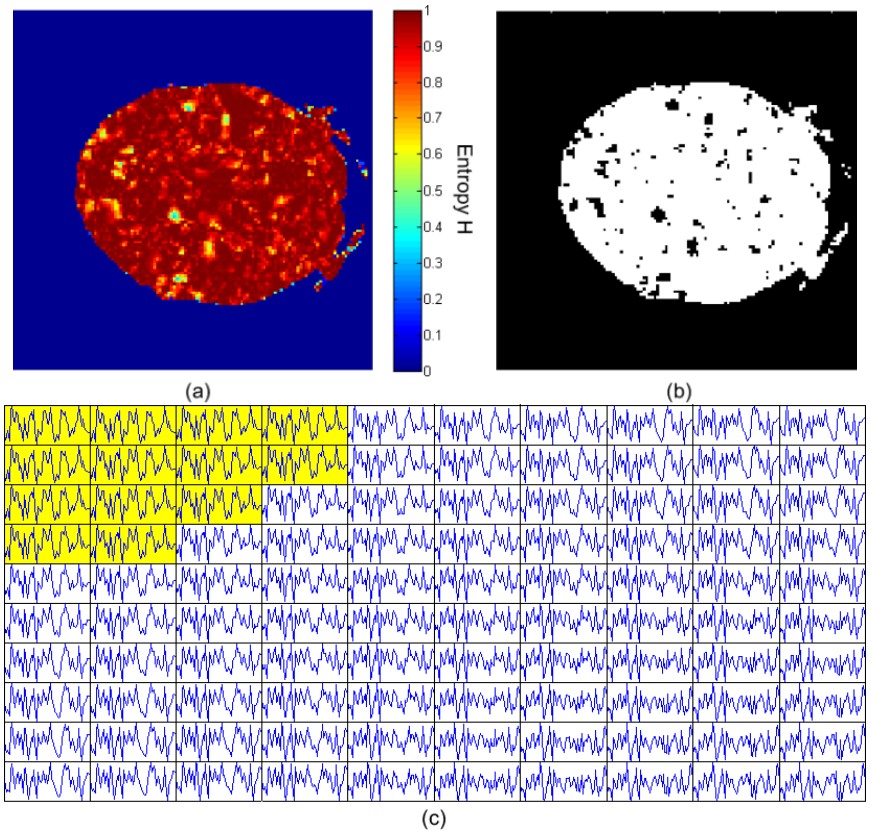


Figure 5. (a) Final configuration of the SOM for the synthetic data, the neurons in highlighted represent the patterns of activation of the input data. (b) Evolution of the quantization error after 250 iterations.



(a) (b)

Figure 6. Result of Simulation (a) Region considered active corresponding to the signal A. (b) Region considered active corresponding to the signal B.



(a) (b)

(c)

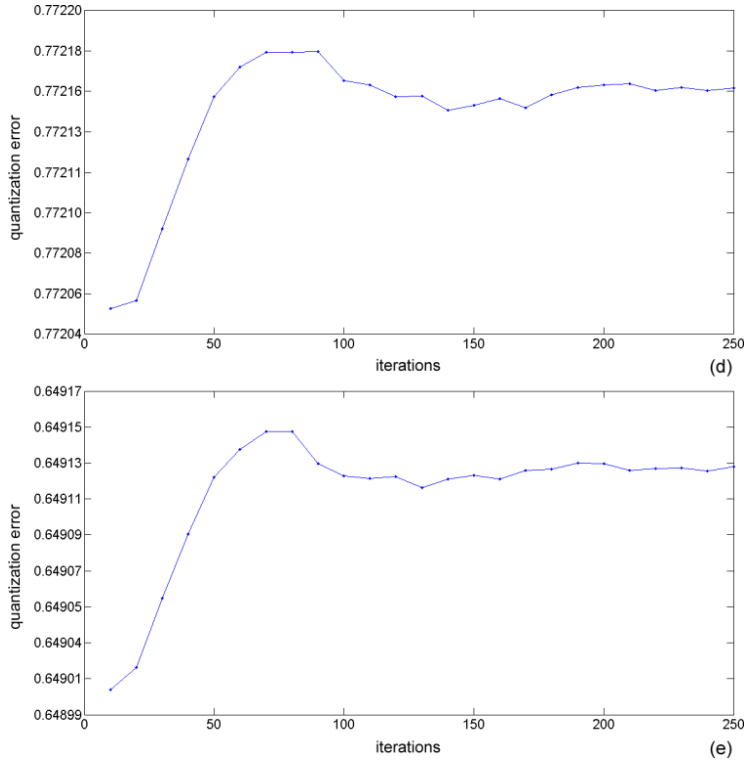


Figure 7. (a) Map of entropy of the real data of the eighth slice. (b) after analysis of entropy, the black dots in the image were eliminated in the subsequent stage, which corresponds to 8% of voxels belonging gray matter. (c) final configuration of the SOM to the real data, the group of neurons highlighted represent the patterns active of the input data. (d) Evolution of the quantization error every 10 iterations without application of entropy (e) Error quantization with entropy.

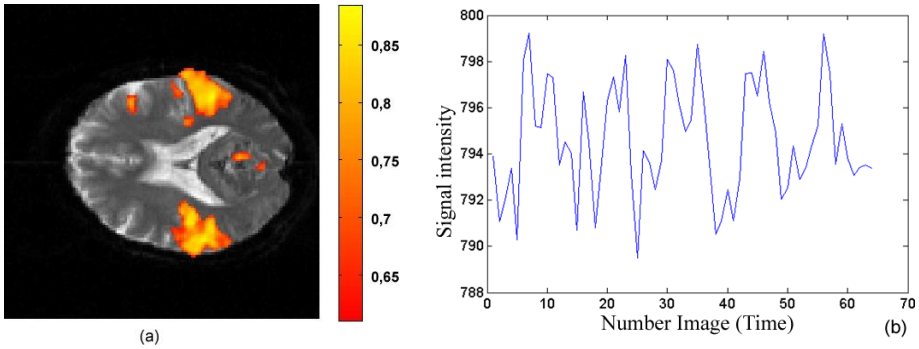


Figure 8. (a) Result of auditory stimulation in the cerebral cortex according to analysis of the SOM, slice 8 (b) Average of the time series of the weights of active neurons.

The activation map constructed from the coefficient of correlation between average weights of the neurons and the time series of the interior of each voxel in brain images can be seen in Figure 8a. In which values with a correlation coefficient greater than 0.8 were highlighted in the brain map. On this map you can see two main regions, as a result of the auditory task, located in the temporal lobe. According to Huettel et al. [20], the temporal lobe also associated with auditory stimuli, may also be related with memory and speech. The Figures 8b show the average weights of the neurons with the application of the entropy.

Figure 9 shows a comparison of the method presented in this article (Figure 9a) with other techniques found in the literature, such as GLM (Figure 9b); Multgrid Priors of Amaral et al. [25] (Figure 9c), and the entropic methods presented by Araujo et al [3] with the entropy of Tsallis (Figure 9d) and of Shannon (Figure 9e).
It can be observed that the result of the procedure presented in this article shows good performance when compared with the GLM and Multgrid priors, the responses are similar, but the active regions demarcated by the SOM are slightly larger than the areas delimited by the GLM and the Multigrid priors. The entropic methods do not work well on auditory data with the block paradigm analyzed here.

4. CONCLUSION

In the present study was used as a tool for data analysis FMRI the self-organizing map of Kohonen, a type of artificial neural network where the training is based on the competitive learning and it makes the topological organization of patterns of input data in a map discrete. Also, with the aim of separating the patterns of the map it was applied hierarchical clustering. Furthermore, in order to increase the efficiency of the method of analysis, it was proposed to use Shannon's entropy to eliminate a range of 5-10% of the set of input data, which in turn were related to voxels that presented signal distributed unevenly.
As can be seen in simulated data, the entropy related to time series of voxels containing evidence of the bold response tends to be maximum. Therefore, the configuration of the data after the entropic analysis allows a better chance of finding groups of active neurons in the grid of the SOM with fewer iteration than the SOM without the entropic analysis. That is, the application of Shannon's entropy to eliminate some voxels improves the performance of the SOM, as it stabilizes the quantization error faster than the SOM without the use of entropy, however, in the data analyzed this improvement was not significant. As compared to other methods for analyzing fMRI data, the procedure presented in this work was as good as the GLM methods and multigrid priors and excellent when compared to entropic methods. It is suggested apply this method to analyse data of FMRI with event related paradigm, it is believed that it will have the same performance that the data with the block paradigm.

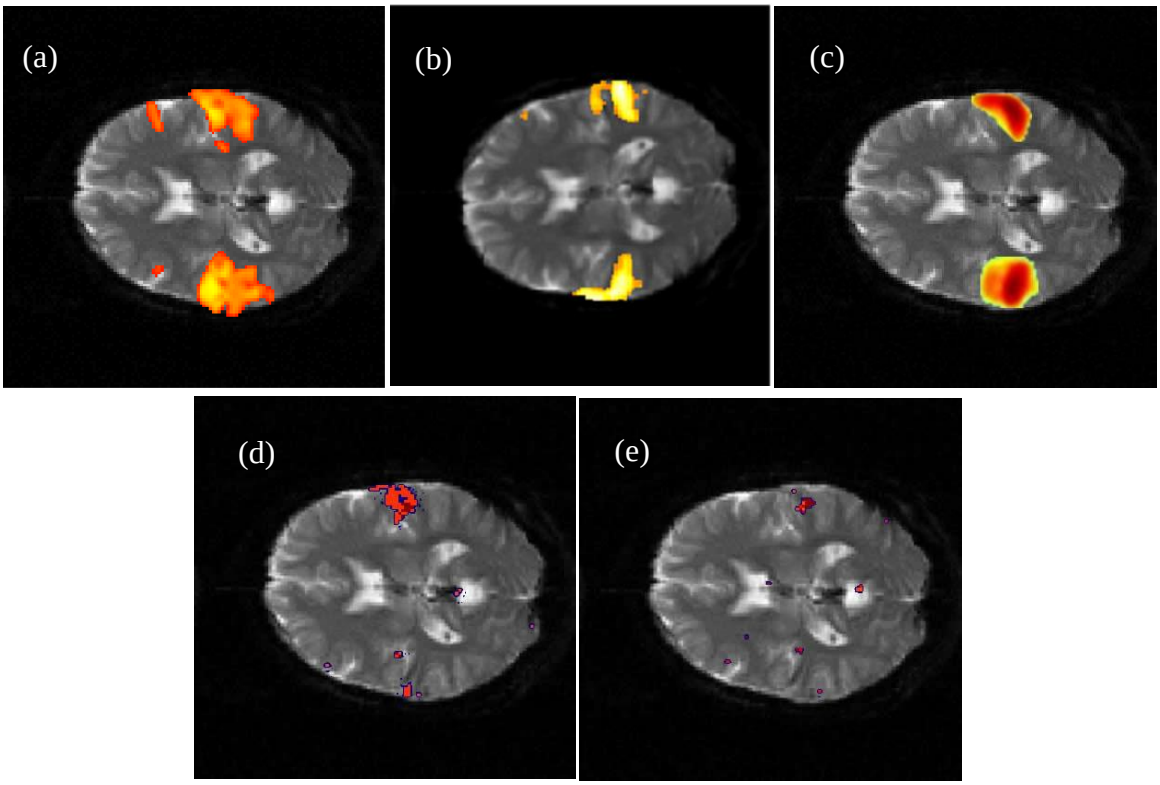


Figure 9. A comparison of the method presented in this article with other techniques found in the literature, the analyses were made on the auditory data with block paradigm (slice 9). (a) Procedure presented in this paper; (b) GLM method; (c) Multigrid priors method; (d) entropic method: Tsallis entropy; (e) entropic method: Shannon entropy.

ACKNOWLEDGMENT.
THIS WORK WAS SUPPORTED BY FEDERAL UNIVERSITY OF PARÁ AND FAPESPA.

REFERENCES

[1] B. B. Biswal, and J. L. Ulmer, Blind source separation of multiple signal sources of fMRI data sets using independent component analysis. Journal of Computer Assisted Tomography, 23(1999), 265-271.

[2] D. Balslev, F. A. Nielsen, S. A. Frutiger, J. J. Sidtis, T. B. Christiansen, C. Svarer, S. R. Strother, D. A. Rottenberg, L. K. Hansen, O. B. Paulson, and I. Law, Cluster analysis of activity-time series in motor learning.Human Brain Mapping, 15(2002), 135-145.
[3] D. B. de Araujo, W. Tedeschi , A. C. Santos, J J. R. Elias, U. P. Neves, O. Baffa, Shannon entropy applied to the analysis of event-related fMRI time series. Neuroimage. 2003 Sep;20(2003):311-7.
[4] G. H. Glover, Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage, 9 (1999), 416-429.
[5] H. Chen, H. Yuan, D. Yao, L. Chen, and W. Chen, An integrated neighborhood correlation and hierarchical clustering approach of functional MRI, IEEE Transactions on Biomedical Engineering, 53(2006), 452-458, 2006.
[6] H. Fischer and J. Henning, Neural-network based analysis of MR time series. Magnetic Resonance in Medicine 41(1999), 124-131.
[7] K. J. Friston, A. P. Holmes, K. J. Worsley, J. P. Poline, C. D. Frith, and R. S. J. Frackowiak, Statistical parametric maps in functional imaging: A general linear approach. Hum. Brain Mapping, 2(1995), 189-210.
[8] K. J. Friston, S.C.R.williams, R. Howard, R.S.J. frackowiak, and R. Turner, Movement-related e ects in fMRI time series. Magnetic Resonance in Medicine, 35(1996), 346-55.
[9] K. J. Friston, P. Fletcher, O. Josephs, A. Holmes, M. D. Rugg, and R. Turner, Event-related fMRI: Characterizing di erential responses. NeuroImage, 7(1998), 30-40.
[10]K. J. Friston, A. Mechelli, R. Turne, and C. J. Price, Nonlinear responses in fMRI: the balloon model, volterra kernel, and other hemodynamics, NeuroImage, 12 (2000), 466-477.
[11] K. J. Friston, Bayesian Estimation of dynamical systems: an application to fMRI, NeuroImage, 16 (2002), 513-530.
[12] K. R. Thulborn, J. C. Waterton, P. M. Matthews, and G. K. Radda, Oxygenation dependence of the tranverse relaxation time of water protons in whole blood at high field. Biochimica et Biophysica Acta, 714(1982), 265-270.
[13] L. Pauling, and C. Coryell, The magnetic properties and structure of hemoglobin, oxyhemoglobin, and carbon monoxyhemoglobin. Proceedings of the National Academy of Sciences of USA, 22(1936), 210-216.
[14] L. Stanberry, R. Nandy, and D. Cordes, Cluster analysis of FMRI data using dendrogram shaperning. Human Brain Mapping, 20(2003), 201-219.
[15]M. Havlicek, K. J. Friston, J. Jan, M. Brazdil, V. D. Calhoun, Dynamic modeling of neuronal responses in fMRI using cubature Kalman filtering, NeuroImage, 56 (2011), 2109-2128.

[16] M. J. Mckeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, Analysis of fMRI data by blind separation into independent spatial components. Human Brain Mapping, 6(1998),160-168.
[17] R. Baumgartner, C. Windischberger, and E. Moser, Quantification in functional magnetic resonance imaging: Fuzzy clustering vs. correlation analysis. Magnetic Resonance Imaging, 16(1998), 115-125.
[18] R. D. Fountura, D. M. Branco, M. Anés, J. C. da Costa, and M. W. Portuguez, Identification of Brain Regions Language: Functional Magnetic Resonance Imaging Study in Patients with Refractory Epilepsy of Temporal Lobe. Journal of Epilepsy and Clinical Neurophysiology, 14(2008):7-10.
[19] R. D. Gibbons, N. A. Lazar, D. K. Bhaumik, S. L. Sclove, H. Y. Chen, K. R. Thulborn, J. A. Sweeney, K. Hur, and D. Patterson, Estimation and classification of fMRI hemodynamic response patterns. NeuroImage, 22(2004), 804-814.
[20] S. A. Huettel, A. W. Song, and G. Mccarthy, Functional Magnetic Resonance Imaging. Sinauer Associates, 1 ed (2004).
[21] S. C. Ngan, E. S. Yacoub, W. F. Auffernann, and X. Hu, Node merging in Kohonen's self-organizing mapping of fMRI data. Artificial Intelligence in Medicine, 25(2002), 19-33.
[22] S. J. Peltier, A. Polk T,D. C. Noll, Detecting low-frequency functional connectivity in fMRI using a self-organizing map (SOM) algorithm. Hum. Brain Mapp., 20(2003), 220-226.
[23] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proceedings of the National Academy of Sciences of USA, 87(1990), 9868-9872.
[24] S. Rabe-hesketh, E. Bullmore, and M. Brammer, The analysis of functional magnetic resonance images. Stat. Methods Med. Res. 6(1997), 215-237.
[25] S. R. Amaral, S. R. Rabbani, and N. Caticha, Multigrid priors for a Bayesian approach to fMRI. Neuroimage 23(2004), 654-662.
[26] T. Kohonen, Self-Organizing Maps. 3 ed. Springer: Berlim, 2001.
[27]Z. Hu, X. Zhao, H. Liu, and P. Shi, Nonlinear analysis of the BOLD signal, EURASIP Journal on Advances in Signal Processing, vol. 2009(2009), 13 pages.
[28] W. Liao, H. Chen, Q. Yang, X. Lei, Analysis of fMRI Data Using Improved Self-Organizing Mapping and Spatio-Temporal Metric Hierarchical Clustering. IEEE Transactions on Medical Imaging, 27(2008), 1472-1483.

Received: October 3, 2013