Proceedings of the 5th International Conference on Information Technology and Application in Biomedicine, in conjunction with
The 2nd International Symposium & Summer School on Biomedical and Health Engineering
Shenzhen, China, May 30-31, 2008

Identifying Eye Blinks in EEG Signal Analysis

Miguel A. Sovierzoski, Fernanda I. M. Argoud, and Fernando M. de Azevedo

***Abstract*—The electroencephalogram is a noninvasive record of the brain activity using electrodes placed on the scalp. The electroencephalographic signal can be contaminated by other signal sources, called artifacts. Among the several artifact sources, eye blink is one of the main sources of interference in the EEG exam, and can be erroneously interpreted as epileptiform activity. This study analyzed eye blink signals acquired by EEG electrodes. The main objective of this study was to develop a neural network classifier specialized in the identification of eye blinks in EEG signals. The statistical study of the eye blinks in EEG signals, the methodology and the results of the identification of this event are presented.**

I. Introduction

THE electroencephalogram (EEG) is the most appropriate exam to analyze neurological pathologies. In electroencephalogram exams, the electrical potential of eye blinks, due to the bioimpedance of the scalp, is acquired along with the brain rhythm signals by the electrodes. This contamination can occur in most electrodes, although it shows a higher intensity in the frontal electrodes. In this case, on the EEG exam, an eye blink is an artifact, which possesses higher amplitude than the brain rhythms. Therefore, in the analysis of the EEG exam it can be erroneously interpreted as an epileptiform activity, which is a characteristic event of epilepsy.

Several studies have been developed with the aim of extracting eye blink artifacts from EEG signals. Reviews and comparisons of the methods to extract or remove eye blinks are presented in the works of Croft and Barray [1], Fatourechi *et al*. [2] and Romero *et al*. [3].

The main objective of this work was to analyze the electrical behavior of eye blink events acquired by EEG electrodes, and also to develop a neural network classifier to identify them. This eye blink event classifier will be used as a part of a hybrid classifier of epileptiform events.

This paper is organized as follows. In Section II, we briefly review the eye blink, Higher Order Statistics, the Multi Layer Perceptron, the standard error backpropagation supervised training algorithm, Cross Validation and ROC Analysis. The results obtained are reported in Section III, followed by concluding remarks in Section IV.

Manuscript received February 29, 2008.

M. A. Sovierzoski is with the Department of Electronics at the Federal University of Technology – Paraná (DAELN-UTFPR) and Biomedical Engineering Institute at the Federal University of Santa Catarina (IEB-UFSC), DAELN-UTFPR, Av. Sete de Setembro, 3165, Curitiba, Paraná, 80230-901, Brazil (www.utfpr.edu.br, e-mail: miguelaso@utfpr.edu.br).

F. I. M. Argoud is with the Federal Center of Technological Education – Santa Catarina (CEFET-SC) and Biomedical Engineering Institute at the Federal University of Santa Catarina (IEB-UFSC), Brazil (www.cefetsc.edu.br, e-mail: fargoud@cefetsc.edu.br).

F. M de Azevedo is with the Biomedical Engineering Institute at the Federal University of Santa Catarina, Brazil (www.ufsc.br, www.ieb.ufsc.br, e-mail: azevedo@ieb.ufsc.br).

II. Materials & Methods

A. Eye Blink

The spontaneous or involuntary eye blink occurs without external stimuli about ten to twenty times a minute, and its function is to clean, lubricate and oxygenate the cornea. The upper eyelid is moved upwards mainly due to the levator palpebrae superioris with a contribution from Müller's muscle. Downward forces are due to stretching of tendons and ligaments of the orbicularis oculi and levator muscles [4].

The eye blink signal acquired by EEG electrodes is a little different from the same event acquired by electrooculography (EOG), due to the different positioning of the electrodes. The EOG electrodes are positioned above and under the eyes and a reference electrode is placed behind the ear.

B. Higher Order Statistics

The EEG signal is stochastic, and each set of samples is called realizations or sample functions (*x(t)*).

The expectance (μ) is the mean of the realizations and is called first-order central momentum. The second-order central momentum is the variance of the realizations. The square root of the variance is the standard deviation (σ), which measures the spread or dispersion around the mean of the realizations [5].

The skewness (third-order central momentum) characterizes the degree of asymmetry of the signal distribution around its mean [5], and is defined in (1).

$$skewness = E\left\{\left[\frac{x(t)-\mu}{\sigma}\right]^3\right\} \quad (1)$$

The kurtosis, also called fourth-order central momentum, characterizes the relative flatness or peakedness of the signal distribution [5], and is defined in (2), which was modified to refer to a non-Gaussian distribution.

$$kurtosis = E\left\{\left[\frac{x(t)-\mu}{\sigma}\right]^4\right\} \quad (2)$$

C. Artificial Neural Network and Training

A Multi Layer Perceptron (MLP) Artificial Neural Network [6] is structured in several layers composed by identical neurons (as shown in Fig. 1), totally connected between adjacent layers.

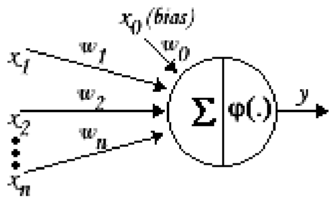


Fig. 1. Structure of a MLP ANN neuron. The bias (x_0) of neuron and every output of the neurons from the previous layer (x_1-x_n) are multiplied by their respective synaptic weight (w_0–w_n) and summed together. The summing represents the total net input of the neuron. This value is passed to an activation function ($\varphi(.)$), to map the continuous output (y) of the neuron.

The standard ANN supervised training algorithm for error backpropagation [6] consists of two steps: the forward propagation and the backpropagation. The forward propagation step is achieved by applying a training pattern to the ANN, propagating it through the network and obtaining the continuous output value. This output value is compared to the desired value of the pattern, generating an error value. The error value is backpropagated to adjust the synaptic weights of the neurons, characterizing the backpropagation step.

The outputs of the ANN are the output of neurons of the output layer, and therefore, assume continuous values, depending on the image of the activation function. However, in most applications, mainly in classifiers, the output values are discretized by a hard limiter function controlled by a threshold, in order to obtain the Boolean values of those outputs.

D. Cross Validation

The Cross Validation (CV) procedure [6], applied to the supervised training of neural networks, evaluates the training and the learning of the ANN. The CV is executed during the ANN training at the end of a training epoch and requires two pattern sets: the training set and the validation set. All training and validation patterns are presented to evaluate the training error and learning error of ANN for that epoch. The errors can be evaluated by the mean square error.

If the training algorithm is converging, the training error is falling towards zero. Normally, the learning error falls to the best generalization point, and then continuously increases, which indicates over-training and the loss of generalization.

E. Statistical Indexes

The comparison between the classification result of the classifier and the indications to a given pattern is made with the contingency table 2x2, resulting in four possibilities: true positive (TP), true negative (TN), false positive (FP) and false negative (FN) [7].

The true indexes (TP and TN) represent the finding that the classifier successfully identified the pattern. The false indexes (FP and FN) represent the finding that the classifier erroneously identified the pattern. These indexes are used to calculate the sensibility and specificity of the classifier [7].

Sensibility is the ability of the classifier to identify the positive patterns among the truly positive patterns [9],[10], and is shown in (3), where (e) is the epoch of training and (t) is the threshold of classifier.

$$sens(e,t) = \frac{\sum TP(e,t)}{\sum TP(e,t) + \sum FN(e,t)} \tag{3}$$

Specificity is the ability of the classifier to identify the negative patterns among the truly negative patterns [9],[10], and is shown in (4).

$$spec(e,t) = \frac{\sum TN(e,t)}{\sum TN(e,t) + \sum FP(e,t)} \tag{4}$$

The values for sensibility and specificity of the neural network classifier are dependent on the training epoch and the threshold value.

F. ROC Analysis

The ROC Analysis evaluates quantitatively the capability of the classifier to perform a classification. The ROC Curve is a Cartesian graph that presents the dependency of sensibility and specificity of a classifier system, and it is independent of the threshold [8], [9]. There are several indexes that can be extracted from the ROC Curve, some of the most important of which include AUC, EER and the SSO index.

The Area Under the ROC Curve (AUC) index assumes values between 0.5, which represents a system unable to discriminate (the number of false indications is greater than the number of true indications), and 1.0, representing an ideal classifier system (no false indications).

The Error Equal Rate (EER) index is obtained at the point of the ROC Curve where sensibility and specificity assume the same value.

The Sensibility to Specificity equal One (SSO) index is taken as the highest sensibility value to specificity equal one, usually at high threshold values.

III. Results and Discussion

A. Statistical Indexes

The realizations of eye blinks are shown in Fig. 2a. Fig. 2b presents the promediation of these realizations, the mean and the standard deviation. An average amplitude of approximately 170 μV is seen for the eye blinks, which is a much higher value in comparison with the brain rhythms, although such an amplitude is compatible with epileptiform activity.

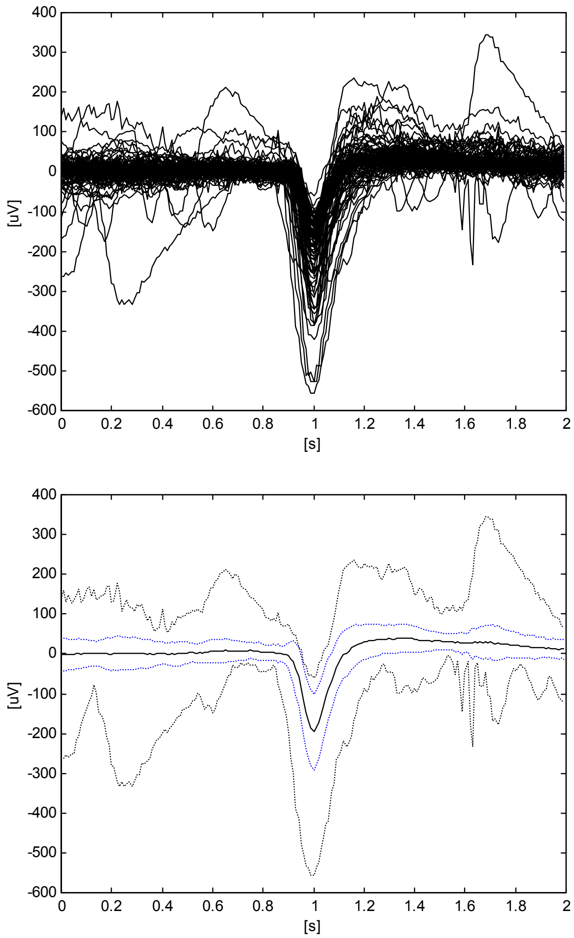


Fig. 2. Graph of the 200 realizations of spontaneous eye blinks acquired with EEG electrodes (a) and graph of the promediation of the 200 realizations of eye blinks (b), showing the superior and inferior masks and the standard deviation. The eyes are opened, then the eyes close (falling curve) and open again (rising curve).

Due to the stochastic characteristic of the EEG signals, there is a large variability in the signals on the realizations graph (Fig. 2a). At the promediation of the eye blinks (Fig. 2b), the amplitude of the EEG signal data is not significant in comparison with the eye blink event.

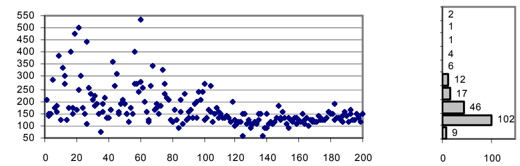
Statistical calculations were made from the eye blink realizations. The behavior of the distribution and the statistical data are presented in the graphs in Fig. 3. Table I shows the summary of the measurements.

TABLE I

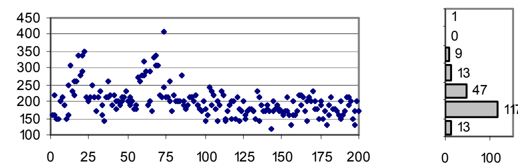
STATISTICS OF 200 REALIZATIONS OF EYE BLINKS

measure	min	$\mu - \sigma$	μ	$\mu + \sigma$	max
amplitude [μV]	55.664	91.723	168.842	245.961	533.202
duration [ms]	120.000	154.623	200.300	245.976	410.000
$t_{closing}$ [ms]	50.000	67.119	85.700	104.280	170.000
$t_{opening}$ [ms]	50.000	79.457	114.600	149.742	290.000
skewness	-0.5678	-0.2940	-0.0883	0.1174	0.5554
kurtosis	1.5261	1.6313	1.7768	1.9223	2.4875

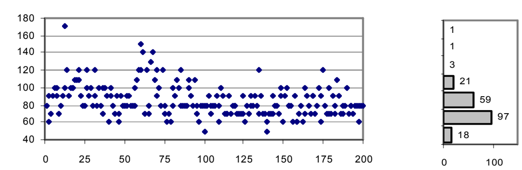
Most values for the amplitude of eye blinks lie in the range of 60 to 300 μV (170 μV mean). The duration varies from 120 to 350 ms (200 ms mean). The spontaneous eyelid closing varies from 60 to 120 ms (85.7 ms mean), while the spontaneous eyelid opening varies from 50 to 175 ms (115 ms mean). The skewness of the eye blink signal varies from -0.5 to 0.4 (-0.09 mean). The kurtosis of the eye blink signal varies from 1.5 to 2.1 (1.78 mean).



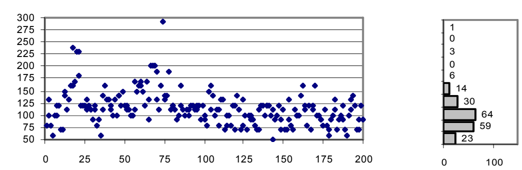
(a) amplitude in [μV] x realization



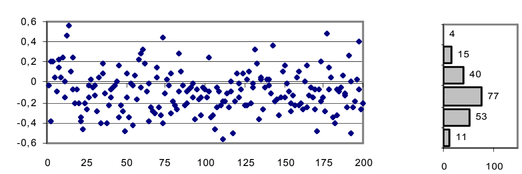
(b) duration in [ms] x realization



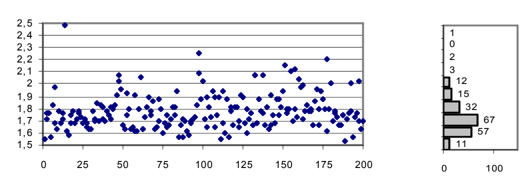
(c) eyelid closing time in [ms] x realization



(d) eyelid opening time in [ms] x realization



(e) skewness x realization



(f) kurtosis x realization

Fig. 3. Measurements and relative distribution graphs of the 200 realizations of eye blinks in relation to amplitude (a), duration (b), eyelid closing time (c), eyelid opening time (d), skewness of the eye blink signal (e) and kurtosis of the eye blink signal (f).

B. Neural Network Classifier

Eye blinks (64 realizations), spikes (76 realizations), noise and normal activity (24 realizations) were separated into a training pattern set (82 events) and a validation set (82 events) of a neural network classifier.

An ANN MLP with three layers and 101:10:1 topology was trained to identify eye blinks in raw EEG data. After dozens of trainings and performance evaluations, the classifier that showed the best results, presented in Fig. 4 (Cross Validation), and performance according to the ROC Curve (Fig. 5), was chosen.

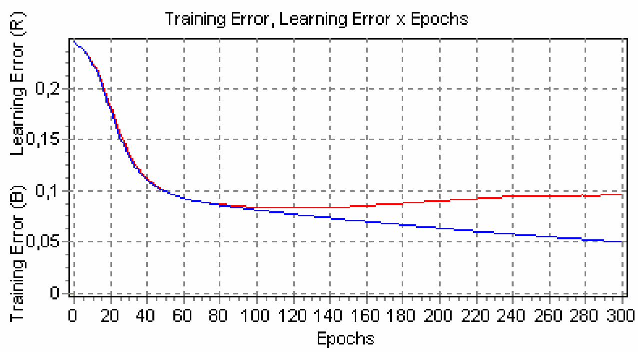


Fig. 4. Graph of the Cross Validation of the training of the neural network classifier. The training error curve is decreasing continuously, which indicates the convergence of the ANN training algorithm. The learning error curve decreases, which indicates the learning phase. At some point, at epoch 120, it starts to rise, indicating overtraining. The best learning level (learning error = 0.0835) was obtained exactly at epoch 120.

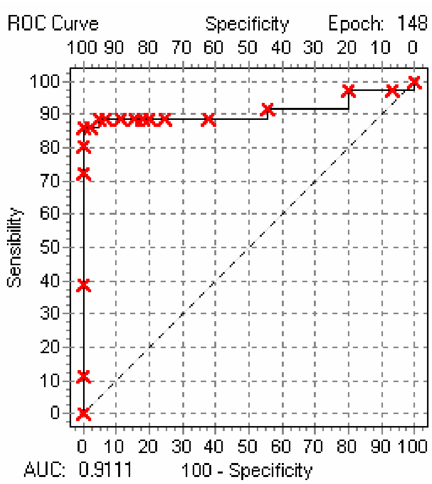


Fig. 5. ROC Curve, obtained at epoch 148 of the eye blink classifier training.

Analyzing the ROC Curve at epoch 148, the following performance indexes were obtained: AUC = 0.9111, EER index at sensibility 88.9% and specificity 95.6%, and SSO index at sensibility 86.1% to specificity = 100%.

By definition, the ERR index should present the same values for sensibility and specificity; however, on analyzing the region close to the crossing point, there is a point with the same sensibility (88.9%) and a greater specificity (95.6%), which was chosen to be the real EER index.

IV. Conclusion

It is important to remember that only spontaneous eye blinks were investigated. Epileptiform events possess wider variability across statistical and morphology parameters, in comparison with spontaneous eye blinks. It maybe that, due to the small variability, the results for sensibility and specificity of the eye blink classifier are rather elevated.

The performance results obtained for the eye blink classifier in this work were obtained in a first analysis. In order to include this classifier as part of a hybrid classifier of epileptiform events, the number of patterns of both training and evaluation sets should be increased, to improve evaluation of the performance, and it is possible that new trainings would be necessary.

References

[1] R. J. Croft, R. J. Barray, "Removal of ocular artifact from the EEG: a review", Neurophysiol. Clin., vol. 30, 2000, pp. 5–19.

[2] M. Fatourechi, A. Bashashati, R. K. Ward, G. E. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey", Clinical Neurophysiology, vol. 118, 2007, pp. 480–494.

[3] S. Romero, M. A. Mañanas, M. J. Barbanoj, "A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: A simulation case", Computers in Biology and Medicine, vol. 38, 2008, pp. 348–360.

[4] M. Iwasaki, C. Kellinghaus, A. V. Alexopoulos, R. C. Burgess, A. N. Kumar, Y. H. Han, H. O. Lüders, R. J. Leigh, "Effects of eyelid closure, blinks, and eye movements on the electroencephalogram", Clinical Neurophysiology, vol. 116, 2005, pp. 878–885.

[5] D. G. Manolakis, V. K. Ingle, S. M. Kogon, Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing, Artech House Publishers, 2005. ch. 3, pp. 75–82.

[6] S. Haykin, Neural Networks: A Comprehensive Foundation, 1998, Prentice Hall.

[7] R. G. Knapp, M. C. Miller, Clinical Epidemiology and Biostatistics, 1992, Harwal Pub Co.

[8] R. C. Eberhart, R. W. Dobbins, Neural Network PC Tools: A Practical Guide, 1990, Academic Press.

[9] T. Fawcett, "An introduction to ROC analysis", Pattern Recognition Letters, vol. 27, 2006, pp. 861–874.

[10] M. A. Mazurowski, P. A. Habas, J. M. Zurada, J. Y. Lo, J. A. Baker, G. D. Tourassi, "Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance", Neural Networks, vol. 21, 2008, pp 427-436.