ARTIFACTS ALLOCATION IN EEG-SIGNALS ON THE BASIS OF DYNAMIC MODELS

Dr. Al-Kasasbeh Riad Taha and Dr. Yousif Eltous
Assistant Professors, Department of Electrical Engineering, Technology Engineering Faculty of Amman, Al-Balqa' Applied University, Jordan .

Two algorithms of artifacts allocation in EEG-signals are described. For signal representation an autoregressive model is used. As a measure of distinction of segments in an EEG-signal the Kullback information is used. Results of real EEG-signal processing are demonstrated.

INTRODUCTION

Due to wide use of digital electroencephalography (EEG) systems in modern medical practice (E. Niedermeyer, F. Lopes da Silva, et al. [1]; F.Lopes da Silva et al. [2]), the problem of automatic allocation of artifacts in EEG becomes important (M.Van de Velde, G. Van Erp, P.J Cluitmans [3]; M.Van de Velde, I.R. Ghosh, P.J. Cluitmans [4]). In existing digital EEG-systems this problem is usually solved by visual analysis of EEG-signal on a screen and its eye-witnesed editing.

In the article two methods of artifacts allocation in EEG-signals are considered. EEG-data is considered as piece-stationary signal. From the random process theory point of view it is possible to consider an artifact as a change in dynamic properties of the process at some moment of time. EEG-signal is presented as an autoregressive process; a set of parameters corresponds to each stationary segment: coefficients of autoregression and dispersion of stimulating white noise.

The first method uses so-called "one-model approach": the first autoregressive (AR) model of necessary order in a fixed window is built on EEG data. It is used as the basic model. Then the second autoregressive model in some sliding window is identified. When spectral measure of distinction between two of these models becomes more than certain threshold, a new segment is formed. After that the logic analysis of fitting to a certain class of artifacts is made. Such procedure of segmentation that uses only a spectral measure of error goes with the delay after the change of spectral density function . As artifacts duration is small, for exact allocation of an artifact in data it is necessary to reduce this delay. For this purpose it is necessary to define the second threshold, smaller than the first one. If the value of the spectral measure of an error exceeds the first threshold then the following stationary segment begins until

the spectral measure of an error becomes equal to the second threshold. Kullback divergence was used as a measure of distinction.

The second method uses so-called "two-model approach", when one model - "global" - is basic, and the other - "local" - is checked. Borders of stationary segments are determined when distinction of parameters of models is essential. Two AR-models are simultaneously estimated: "local" – is estimated in short temporary window, corresponding to a signal after probable change, and "global" – in long temporary window, corresponding to a signal before change.

AIM AND METHOD OF THE RESEARCH

Description of One- Model Algorithm

Segmentation of EEG-data is based on stationarity analysis of value of a spectral measure of error, which represents distinction between two spectral functions of density.

Basic model is an autoregressive model in a fixed window 0, the second model - in some sliding window M'-M. Allocation of windows is shown in figure 1.

EEG-data was considered as consisting from stationary subsequences. That means, the spectral estimation has insignificant changes in time. In 1980 H.Sugimoto has developed an adaptive method of segmentation for EEG analysis (N. H. Sugimoto [5]).

We shall define the Kullback information $I(1:2)$ as:

$$I(1:2) = \int f_1\left(x_1, \Lambda, x_k\right) \ln \frac{f_1\left(x_1, \Lambda, x_k\right)}{f_2\left(x_1, \Lambda, x_k\right)} dx_1 dx_2 \Lambda \ dx_k$$

Where $f_1(x_1, \dots, x_k)$ and $f_2(x_1, \dots, x_k)$ are functions of density of distribution of 1st and 2nd class accordingly. Provided that $f_1(x_1, \dots, x_k)$ and $f_2(x_1, \dots, x_k)$ are normal density of distribution of $N(\mu_1, \Sigma_1)$ and $N(\mu_2, \Sigma_2)$ accordingly, the Kullback information becomes:

$$I(1:2) = \frac{1}{2}\ln\frac{|\Sigma_2|}{|\Sigma_1|} + \frac{1}{2}tr\Sigma_1(\Sigma_2^{-1} - \Sigma_1^{-1}) + \frac{1}{2}tr\Sigma_2^{-1}(\mu_1 - \mu_2)(\mu_1 - \mu_2)$$

Where μ_1 and μ_2 - mean values and Σ_1, Σ_2 - covariance matrices accordingly.

Kullback information that is limited in time is defined as function from length of data set T:

$$I_T(1:2) = \frac{1}{2}T^{-1}\left[\frac{|\Sigma_2|}{|\Sigma_1|} + tr\Sigma_1\Sigma_2^{-1} - T\right] + \frac{1}{2}T^{-1}\delta^T\Sigma_2^{-1}\delta$$

Where
$$\mu_i = \left[\mu_i(0), \dots, \mu_i(T-1)\right]$$
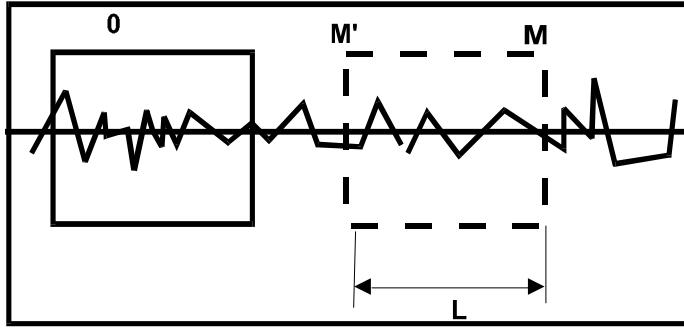$$\Sigma_i = \{\sigma_i(s-t); \ s, t = 0, 1, \dots, T-1\}$$
$$\delta = \mu_1 - \mu_2$$

Fig. 1: Windows in one-model algorithm.

If length T is sufficiently great, then:

$$\lim_{T \to \infty} T^{-1} \log | \Sigma_i | = \int_{-\pi}^{\pi} \log p_i \, (w) \frac{dw}{2\pi}$$

It is proved (A. Afifi, S. Azen [6]) that Kullback information limited in time $I_T \, (1:2)$ aspires to $I(1:2)$ at rather large T. We shall define spectral density $P_1(\omega)$ as

$$P_1(\omega) = \frac{\sigma_1^2}{\left| A_1 \big[\exp \, (j\omega) \big] \right|^2}$$

Other density $P_1(\omega)$ is accepted as

$$P_2(\omega) = \frac{\left| E_2 \big[\exp \, (j\omega) \big] \right|^2}{\left| A_2 \big[\exp \, (j\omega) \big] \right|^2} = \frac{\left| E_2 \big[\exp \, (j\omega) \big] \right|^2}{\left| A_1 \big[\exp \, (j\omega) \big] \right|^2}$$

It is here accepted that $A_1 \big[\exp \, (jw) \big] = A_2 \big[\exp \, (jw) \big]$

It is possible to write down the following equality:

$$\frac{P_2(\omega)}{P_1(\omega)} \sigma_1^2 = \gamma_0 + \sum_{k=1}^{m} \gamma_k \cos (\omega k)$$

Where $\{ \gamma_k \}$ $(k = 0, 1, \Lambda, m)$ is autocorrelation function of a residual sequence of autoregressive model (S.L. Marple [7]). Applying mentioned formula without constant coefficients we shall present in a following way:

$$MCO = (\sigma^2 - \gamma_0) + 2 \sum_{k=1}^{m} \gamma_k^2$$

Here MCO is spectral measure of an error, first term reflects changes of complete energy of signal, while the second term reflects only changes of a spectrum.

If value of a spectral measure of an error exceeds a threshold θ_1, segmentation is carried out. Then, the sequence after segmentation is considered as different from a sequence before segmentation in sense of spectral density.

In general, the segmentation that uses only a spectral measure of an error, goes with a delay after the change in spectral density function. As the duration of artifacts is rather small (less then one second), for exact allocation of artifact in data it is necessary to remove this delay. Hence the second threshold θ_2 is given to reduce the delay in segmentation. The role of these thresholds is shown in the following way. When a value of the spectral measure of an error exceeds the threshold θ_1, the segmentation turns back until a value of a spectral measure of an error becomes equal to the second threshold θ_2 smaller than θ_1.

DESCRIPTION OF THE TWO- MODEL ALGORITHM

The given algorithm realizes so-called "two-model approach" for detection of changes in AR-process. Two AR-models are simultaneously estimated: "local" and "global". Allocation of windows in case of use of such approach is shown in figure 2.

Let (y_n) scalar signal, described by model:

$$y_n = \sum_{i=1}^{p} a_i^{(n)} y_{n-i} + \varepsilon_n$$

$$\text{var}(\varepsilon_n) = \sigma_n^2$$

Where:

$$\left. \begin{array}{c} a_i^{(n)} = a_i^0, \quad 1 \le i \le p \\ \sigma_n^2 = \sigma_0^2 \end{array} \right\} \quad n \le r-1$$

$$\left. \begin{array}{c} a_i^{(n)} = a_i^1, \quad 1 \le i \le p \\ \sigma_n^2 = \sigma_1^2 \end{array} \right\} \quad n \ge r$$

Where (e_n) - white noise, r-a moment of change of a signal.
We shall consider statistic of the following kind

$$W_n = \sum_{k=1}^{n} w_k$$

In which

$$w_n = -\int g^0(y/Y^{n-1}) \ln \frac{g^1(y/Y^{n-1})}{g^0(y/Y^{n-1})} dy + \ln \frac{g^1(y/Y^{n-1})}{g^0(y/Y^{n-1})}$$

The conditional mathematical expectation before and after the change:

$$E_0(w_n/Y^{n-1}) = 0$$
$$E_1(w_n/Y^{n-1}) = J(g^0, g^1) > 0$$

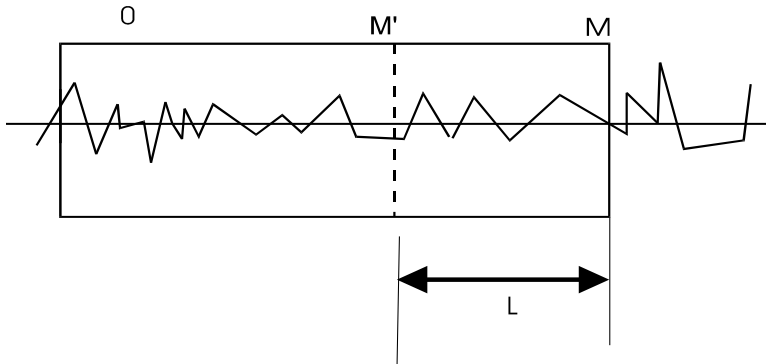Here $J(g^0, g^1)$ is a Kullback divergence.

Fig. 2: Windows in two-model algorithm.

The drift after the change is a symmetric function, that allows to realize symmetric detection, i.e. comparable characteristics of algorithm at transition from a "pure" signal to "noised" and vice versa.

In case of Gauss autoregression w_n is expressed:

$$w_n = 1/2 \left[-2 \frac{e_n^0 e_n^{10}}{\sigma_1^2} + (1 + \frac{\sigma_0^2}{\sigma_1^2}) \frac{\left(e_n^0\right)^2}{\sigma_0^2} - 1 + \frac{\sigma_0^2}{\sigma_1^2} \right]$$

The conditional mathematical expectation before and after the change is:

$$E_0(w_n / Y^{n-1}) = 0$$

$$E_1(w_n / Y^{n-1}) = -1 + \frac{1}{2}\left(\frac{\sigma_0^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_0^2} \right) + \frac{1}{2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_0^2} \right)\left(e_n^1 - e_n^0\right)^2 > 0$$

Here e_n^0, e_n^1 are errors of updating for the first and the second model accordingly.

The Burg algorithm was used as an algorithm that was estimating AR-parameters in a "short" temporary window, and a size of a window L can be changed (various size of a window was tested and good results were achieved at the size in 3-4 times more than order of AR-model). The trellised filter on the basis of approximate algorithm of the least squares was used as algorithm for estimation of AR-parameters in a "long" temporary window (A. Afifi, S. Azen [6]).

The Hinkly test was used to W_n statistics to decrease delay and to obtain good estimation of the moment of change, i.e. the following statistics was used:

$$\widetilde{W}_n = \sum_{k=1}^{n}(w_k - \delta)$$

Where δ is a positive drift.

Typical behaviour of statistics is shown in figure 3.

The moment of n_d disclosure corresponds to crossing of threshold λ by the value:

$$Z_n \equiv \frac{\max}{1 \le m \le n} \tilde{W}_m - \tilde{W}_n$$

Moment of spasmodic change of the model is determined in a following way:

$$r = \sup\left\{ n \big| n \le n_d, \tilde{W}_n = \frac{\max}{1 \le m \le n_d} \tilde{W}_m \right\}$$
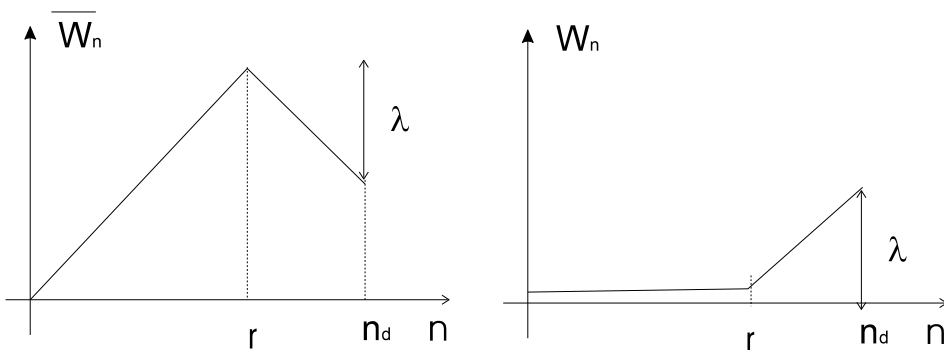

Fig. 3: Typical behaviour of statistics is shown on the 3rd figure.

EXPERIMENTS

For testing the algorithms we used EEG data that was obtained during scheduled diagnostic work in clinical laboratory. EEGs of eleven patients were used. During the experiments patients were in a quiet wakefulness condition (with closed or opened eyes). All patients had no steadily rough infringements of bioelectrical activity of the brain.

Files, containing digital EEG's, were recorded by serial system "Encephalan 131-01, v. 4.2", manufactured by "Medicom LTD " company. Recordings were made under following parameters:

- Number of channels - 16;
- Sampling frequency - 150 Hz;
- Number of digits of analog-digital converter - 10;
- Duration of recording during investigation - up to 15 minutes.

Fragments containing various displays of bioelectrical activity and artifacts were used for EEG processing. Allocation and interpretation of EEG fragments were made by doctor-expert by means of interactive analysis of records.

During analysis of each fragment expert visually allocated parts of record that contained artifacts, parts of record with paroxysmal activity and parts of record

without the specified displays. The necessity of division of parts of records with paroxysmal activity and without it is caused by a fact that paroxysm and artifacts look similar and consequently their division requires additional analysis.

As it was mentioned earlier, artifacts depending on the reason of their occurrence are divided into groups. For unity of data description, groups of paroxysmal and regular activity were included in classification. Thus, depending on kind of distortion the following groups were generated:
–   "Blinking" (group 1);
–   "Eyes movement" (group 2);
–   "Electrode movement" (group 3);
–   "Bad contact" (group 4);
–   "Paroxysm" (group 5);
–   "Regular activity" (group 6).
The following parameters were used:

Two-model algorithm.
–   The order of autoregressive model: 20 (estimation of autoregressive coefficients by Burg algorithm (S.L. Marple [7]));
–   Size of drift: 1.0:
–   Length of local window: 100 samples (0.66 sec):
–   Threshold size: 7;

One-model algorithm.
–   The order of autoregressive model: 20 (estimation of autoregressive coefficients by Burg algorithm);
–   Length of window: 300 samples (2 sec);
–   Size of threshold 1 θ_1 : 6;
–   Size of threshold 2 θ_2 : 4.

In table 1, results of algorithms used are submitted.

CONCLUSION

As one can see from table 1, the both developed algorithms give good coincidence with expert estimations because level of coincidence is no less than 91%. According with artifact structure one model algorithm more sensitivity to detect blinking and electrode movement, two-model algorithm demonstrated better result to detect eyes movement and paroxysm artifact.
In whole, comparing common ability of both artifact detection methods, two-model algorithm gives better results in comparison with one-model algorithm, but it is necessary to take into account, that two-model algorithm has larger computing complexity.
Reasonable agreement of results of automatic allocation with expert estimations allows to recommend to include developed methods in software of modern computer electroencephalographic systems.

Table 1: Results of algorithms use in case of different artifacts in EEG.

Artifact	Number of realization	one-model algorithm		two-model algorithm	
		coincidence with expert	false alarm	coincidence with expert	false alarm
Blinking	33	93.1%	14.5%	92%	9.3%
Eyes movement	28	91%	9.1%	97,5%	6.7%
Electrode movement	24	93.2%	11.4%	95,7%	7.2%
Bad contact	37	91.5%	17.3%	94,1%	9.1%
Paroxysm	31	92.2%	8.7%	98,2%	5.7%

Further development of suggested methods lies in solving of a problem of automatic classification of types of allocated artifacts by means of analysis of characteristics of spectral power density that is estimated on local parts of signal with artifacts (R. T. Al-Kasasbeh [8]). Technology of wavelet transform can be effectively used for this problem solving.

ACKNOWLEDGEMENTS

REFERENCES

[1] E. NIEDERMEYER, F. Lopes da Silva, et al., Electroencephalography: basic principles, clinical applications, and related fields, Baltimore, 1993.

[2] ED. BY F.LOPES DA SILVA ET AL, Handbook of Electroencephalography and Clinical Neurophysiology - V. 2: Clinical applications of computer analysis of EEG and other neurophysiological signals, Amsterdam, 1986.

[3] M.VAN DE VELDE, G. Van Erp, P.J Cluitmans, Detection of muscle artifact in the normal human awake EEG, Electroencephalogr. Clin. Neurophysiol. № 107(2), p. 149-158, 1998.

[4] M.VAN DE VELDE, I.R. Ghosh, P.J. Cluitmans, Context related artifact detection in prolonged EEG recordings, Comput. Methods Programs Biomed. № 60 (3), p. 183-196, 1999.

[5] N.H.SUGIMOTO, Computer classification of the EEG time series by Kullback information measure, INT.J.SYSTEMS SCI. vol. 11, № 6, p. 677-687, 1980.

[6] A.AFIFI, S.AZEN, Statistical Analysis, A computer Oriented Approach, Academic Press. New-York, 1979.

[7] S.L.MARPLE , Digital Spectral Analysis, Prints_Hall, Inc., Englewood Cliffs. N.Jersey, 1987.

[8] R. T. Al-Kasasbeh , Wavelet-based method for EEG artifacts classification. FMBE Proceedings, pp.157 – 159, Kuala Lumpur, Malaysia, 2004.

اختيار الظواهر الإختلالية لإشارات فروق جهود الكهرباء الدماغية (EEG) باستعمال معلوماتِ Kullback

هـذه الورقـة البحثيـة تتطـرق لدراسـة خوارزميتـان تسـتخدمان لتحديـد الظـواهر الإختلاليـة عـن الظـواهر الطبيعيـة **(artifacts)** لإشـارات فـروق جهـود الكهربـاء الدماغيـة **(EEG)**، علـى سـبيل المثال قيـاس الأجـزاء المميـزة مـن إشـارة **(EEG)** مسـتخدمين معلومـات **Kullback** كـل ذلـك يتيـح إعطـاء التشـخيص الطبـي السـليم أخـذت الدراسـة علـى إشـارات واقعيـة حقيقيـة **(EEG)** وأثبتـت فاعليتها في ذلك.حيث يمكن عرض إشـارات فـروق جهـود الكهربـاء الدماغيـة **(EEG)** علـى شـكل نموذج الرياضي **Autoregressive** .

هاتــان الخوارزميتــان تتطــابق نتائجهمــا مــع تقيـيم الخبـراء البـاحثين عـن تلـك الإخـتلالات **(artifacts)**. في ضوء ذلك تسمح لنا بالتوصية بإضافة تلك الخوارزميات علـى شـكل برمجيـات و وربطـه مع مضخمـات **EEG**. لتطـوير ذلـك الإتجـاه الخوارزميتـان يمكـن عـن طريـق نظريـات تحويـل الموجـة و خصـائص الشـبكات العصـبونية لإستخلاص ميزات تلك الإختلالات و الأسـباب المؤديـة لهـا و تصنيفها. إن إئتلافيـة تلك الخوارزميات تجعل إكتشاف الإختلالات الحقيقيـة لإشـارات فروق الجهد الكهربائية **EEG** أكثر فاعلية.

الخوارزميــة الثانيــة تعطــي نتــائج تحديـد الإخـتلالات الناتجــة عــن إزاحــة المجســات **(electrode movement)** و حركـة العـين **(eyes movement)** و مـرض **(paroxysm)** و الـ **blinking** أكثر فاعليـة مـن الخوارزميـة الأولـى و لكـن الخوارزميـة الثانية معقدة رياضيا و تحتاج إلى وقت أطول في **حساب** تلك الإختلالات.