ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 *2007. No. 1(73)*
ELEKTRONIKA IR ELEKTROTECHNIKA

MEDICINE TECHNOLOGY
T 115
MEDICINOS TECHNOLOGIJA

Segmentation of Electroencephalographic Signals using an Optimal Orthogonal Linear Prediction Algorithm

R. Ursulean
Faculty of Electrical Engineering, "Gh.Asachi" Technical University
Bd. D. Mangeron 53, Iasi 700050, Romania; e-mail: ursulean@ee.tuiasi.ro
A. M. Lazar
Faculty of Bioengineering, "Gr. T. Popa" Medicine and Pharmacy University,
Universitatii str. 16, Iasi, 700115, Romania; e-mail:anca@scs.etc.tuiasi.ro

Introduction

The analysis and processing of electroencephalographic (EEG) signals are carried out for many reasons, the most important ones are to implement brain-computer interfaces, to make diagnosis by means of automated classification and to use the features of these signals as biometric keys.

The first stage in all the above-mentioned tasks is to drastically reduce the amount of collected data. This may be achieved by means of evidencing the stationary parts of the signal as the EEG signal has an important inherent feature, a high non-stationarity.

In this paper, we present an adjusted method that serves for the segmentation of the EEG signal into stationary fragments. The mathematical background that describes the algorithm is presented in the first section of the paper, the second section deals with the experimental results, the third is a discussion of the possible applications and the last one is devoted to the conclusions.

Theoretical background

Segmentation and detection of different waves of an EEG signal may be obtained in several ways. Among them is the one based on the so-called "innovations filter", first introduced by I. Schur, [1], [2]. The "innovations filter" is, in its fundamental nature, an optimal orthogonal linear prediction algorithm.

This choice is appropriate when dealing with non-stationary signals with randomly appearance that are of unknown amplitude, which is the case of the EEG signals. Because it involves an adaptive optimal orthogonal parameterization, the algorithm that governs the functioning of the filter is well suited for both detection and segmentation of the signal.

In what follows, we shall use the notations introduced by Lee, Morf and Friedlander in [3], to review the mathematical background.

An "innovations filter" is implemented using N identical sections, each one described by means of a reflection coefficient, ρ, a forward prediction error, ν and backward prediction one, η. The essence of the algorithm is that it calculates the parameters of the model for each new sample of the signal and that explains why the "innovations" term was chosen for the filter.

The recursive equations that define the n-th section of the filter, at a certain t moment, are the following:

$$\rho_{n,t} = \rho_{n,t-1} \cdot \sqrt{(1-\nu_{n-1,t}^2)(1-\eta_{n-1,t}^2)} - \nu_{n-1,t} \cdot \eta_{n-1,t-1}, \quad (1)$$

$$\nu_{n,t} = \frac{\nu_{n-1,t} + \rho_{n,t} \cdot \eta_{n-1,t-1}}{\sqrt{(1-\rho_{n,t}^2)(1-\nu_{n-1,t}^2)}}, \quad (2)$$

$$\eta_{n,t} = \frac{\eta_{n-1,t-1} + \rho_{n,t} \cdot \nu_{n-1,t}}{\sqrt{(1-\rho_{n,t}^2)(1-\nu_{n-1,t}^2)}}, \quad (3)$$

with $n \in \{1,2,...N\}$ and $t \in \{1,2,...M\}$.

It is worth mentioning that the algorithm uses normalized values for both the coefficients and the signal. The signal is normalized by means of its variance, denoted by R_t, as given by the following equation:

$$R_t = \lambda_t \cdot R_{t-1} + x_t^2, \quad (4)$$

where x_t is the sample of the signal at the t instant.

Therefore, the normalized value of the signal is

$$\hat{x}_t = \frac{x_t}{\sqrt{R_t}}. \quad (5)$$

The main advantage of using normalized values lays in numerical stability (all the values used in computation are between 0 and 1).

The normalization process involves also the so-called forgetting factor, $\lambda_t \in [0,1]$. The forgetting factor has the role of weighting the previous value of the signal,

compared to the new sample. This factor may be kept constant but, as it will be shown, it is far more practical when the forgetting factor is computed in relation with the energy of the forward prediction error for the last section of the "innovations filter", $\nu^2_{N,t-1}$, as given by the subsequent formula:

$$\lambda_t = a \cdot \lambda_{t-1} + (1-a) \cdot (1 - \nu^2_{N,t-1}), \quad (6)$$

where $a \in (0,1)$ is a parameter used to adjust the weight of the former value of λ.

The forgetting factor is a good indicator of the stationarity of the signal and it is possible to make use of it for the segmentation of the signal. When its value is relatively constant, it may be considered that the signal is stationary or quasi-stationary; on the other hand, dealing with important variations of the forgetting factor is a clear marker of the presence of non-stationarities in the signal. In order to measure the variations of the forgetting factor, its values have to be compared to a threshold. To obtain an adequate segmentation it is worth noticing that this threshold must not be a constant value.

We suggest the following formula for the threshold of the forgetting factor, λ_{th}:

$$\lambda_{th} = \overline{\lambda} \cdot (1 - 3\sigma_\lambda), \quad (7)$$

where $\overline{\lambda}$ is the mean and σ_λ is the standard deviation of the forgetting factor. We choose this expression because it is closely linked to the statistical properties and therefore the forgetting factor is able to adjust its values according to these properties.

This choice was made taking into account the fact that the EEG signals are of very different shapes and so the segmentation procedure is based on statistical considerations, i.e. stationarity.

Experimental results

The datasets used in our study were downloaded from http://republika.pl/eegspikes, a site created by M. Latka. The 19 channels EEG signals, sampled at 240 Hz, recorded according to the international 10-20 standard system were from juvenile epileptic patients. The signals were divided into three groups, as follows:

1. signals with large, single spikes which are not accompanied by the prominent slow wave (30 epochs, files labeled *s1* to *s30*);
2. signals with spikes or sharp waves followed by slow waves with comparable amplitudes (14 epochs, files labeled *s38* to *s51*);
3. a sequence of spikes (spikes localized close to each other, 7 epochs, files labeled *s31* to *s37*).

The initialization of the three parameters of the innovation filter was chosen as follows:

$$\rho_{n,0} = 0, \quad \nu_{0,0} = x_0, \quad \eta_{0,0} = x_0 . \quad (8)$$

Besides this, the initial values for the forgetting factor and the variance were chosen as

$$\lambda_0 = 1, \quad R_0 = x_0^2 + \varepsilon, \quad (9)$$

where ε (a very small quantity) was introduced only for computational reasons, to be sure that division by zero is avoided, because we cannot know in advance that the value of the first sample of the signal is not zero.

From various runs of the algorithm we found that N=11, representing the number of sections of the filter, was a good compromise between the amount of computing and the segmentation capabilities.

In our work, we also found that setting a = 0.2 in formula (6) is a good choice for weightings of the forgetting factor at time *t-1* and the energy of the forward prediction error of the last section of the filter, at the same moment.

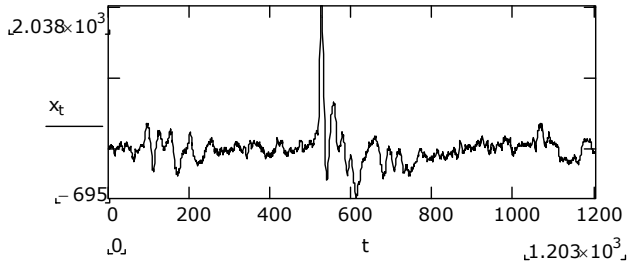

Fig. 1. Signal s10 (epoch 10)

To illustrate the need for a threshold value λ_{th}, linked to the statistical properties of the signal that has to be segmented, let us consider the signal x_t in Fig. 1, the forgetting factor λ_t and its threshold value λ_{th} for the same signal, in Fig. 2.

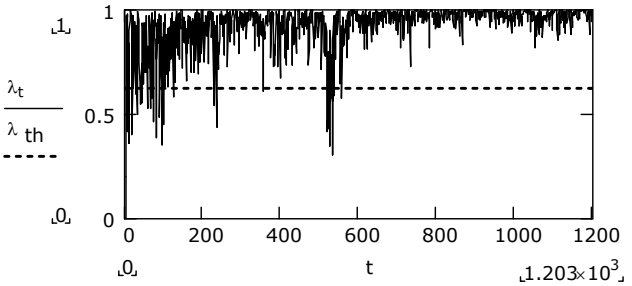

Fig. 2. The forgetting factor λ_t and its threshold value λ_{th} for the signal from Fig. 1

In Fig. 2 there are visible smaller values of λ_t for the first samples, $t \in [0,100]$, of the signal, because the algorithm needs a certain period of time to adapt itself to the signal. In this particular case, the threshold value was computed, according to the previous formula, as λ_{th}=0.618. It is easy to notice the fact that there is a zone (around $t \in [540, 580]$) evidenced by small values of the forgetting factor and about three zones with forgetting factors higher than the threshold (the ending zone is characterized even by values near unity). As values behind the threshold are obvious signs of stationarity, we could say that this signal may be divided into three stationary segments, as it follows: [100, 250], [250-540] and [580-1200].

In fact, the value of the forgetting factor is also a measure of the derivative of the signal: for equal amplitude variations, in different periods, the forgetting factor varies slower in the case of the longer period.

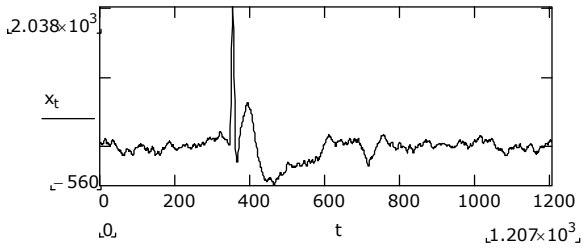

Fig. 3. A single-spike signal

For signals that do not exhibit many deviations, known as "single-spikes", see Fig. 3, a typical detection is as the one presented in Fig. 4.

According to the graph in Fig. 4, there are only two segments of the signal, the first one being obviously non-stationary.

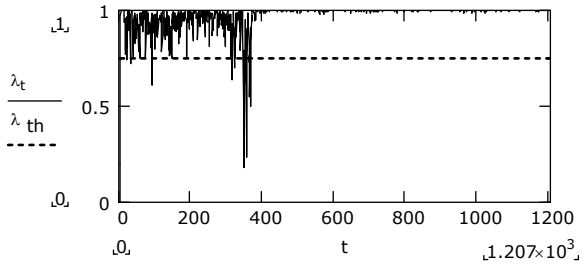

Fig. 4. The forgetting factor and its threshold value for the single-spike signal in Fig. 3

The spike-sequence signal presented in Fig. 5, which is far different from the previous ones can also be segmented by means of proper choosing of the threshold value.

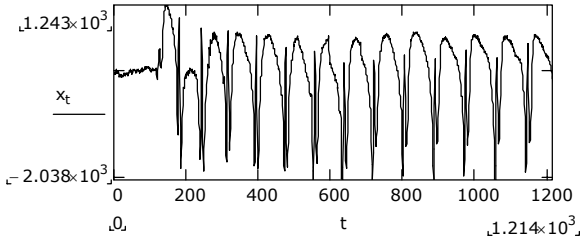

Fig. 5. The spike-sequence signal

This kind of signal is somewhat more difficult to characterize since the forgetting factor, as seen in Fig. 6, exhibits values nearly equal to the threshold value, λ_{th}=0.599, even in the stationary region, for t > 300. Nevertheless, the overall tendency for λ_t is an increasing one and this is clearly seen from the graphic.

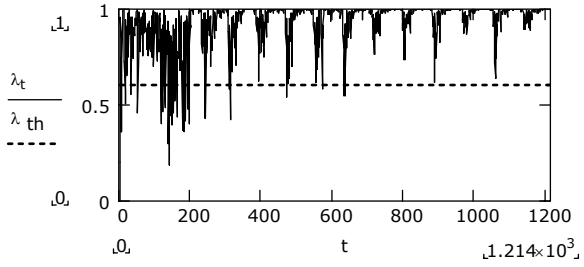

Fig. 6. The forgetting factor and its threshold value for the spike-sequence signal in Fig. 5

The last type of signal that was the subject of the segmentation procedure was the so-called "slow-spikes", presented in Fig.7.

Here it seems that the non-stationarities are more difficult to be evidenced because of the slower variations.

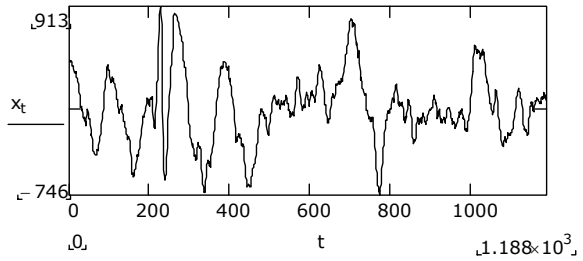

Fig. 7. The slow spike signal

Nevertheless, choosing a threshold linked with the statistical features of the signal, like the one given in equation (7), proved a good choice since there are moments in which the threshold is greater than the forgetting factor, as it can be seen in Fig.8.

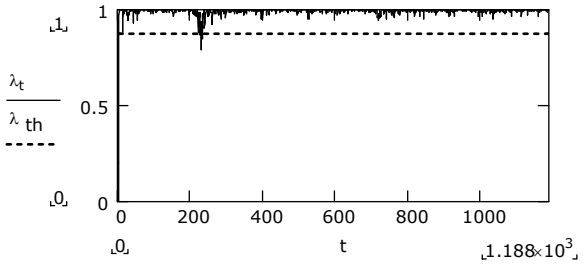

Fig. 8. The forgetting factor and its threshold value for the slow-spike signal

It is apparent from Fig. 8 that a constant threshold could lead to wrong segmentation since the variations are somewhat less significant due to their slow nature. In this case λ_{th}=0.871 and there is just one period of time, around t=250, when this value is exceeded.

Discussion

One of the applications of the above method of segmentation is in predicting the seizures in epilepsy. The capability to predict epileptic seizures prior to their occurrences may end in novel diagnostic means and treatments of epilepsy, [4]. Between seizures, some waves, morphologically defined episodic, transient EEG discharges, were found. Spikes, sharp waves, slow waves and slow-wave complexes are the types included in this category, [5]. A comprehensive study on different methods on seizure prediction may be found in [6]. The authors, using comprehensive references, concluded, based on several methods, that there is strong evidence that seizures can be predicted by means of the interictal EEG signals.

During last years, some approaches centered both on spatial and temporal contextual information in detecting epileptiform activity waves were developed, [5], [7]. These papers proposed multistage systems, some of them involving, besides the well-known stage of feature extraction, a preliminary stage dedicated to separate the

stationary and non-stationary components of the EEG signal. In this way, data volume is significantly reduced.

From this point of view, our work may be of real interest since the segmentation of the signal may be achieved fast and with low computing power.

Another possibility to use the results is the one that arise from the use of EEG signals as biometric keys. In this case, a proper segmentation, according to the stationarity of the signal, may improve significantly the rate of success when authenticating with a system using EEG signals as biometric keys.

Conclusions

We have tailored a method based on an optimal orthogonal linear prediction algorithm to obtain the segmentation of an EEG signal into stationary components. A close link of the threshold of the forgetting factor to the statistical properties of the signal, in spite of the usual constant values used in other applications, was the key to evidence the non-stationary parts of the signal.

We have applied the method on three types of interictal epileptic EEG signals: signals with large, single spikes or sharp waves which are not accompanied by the prominent slow wave, signals with spikes or sharp waves followed by slow waves with comparable amplitudes and sequences of spikes. The best results were pointed out for the first two types of interictal spikes. Nevertheless, when dealing with spike-sequences EEG signals, the algorithm reached its lowest rate of success.

The future work will focus on comparing the qualities of this method to those evidenced by other methods (e.g. the use of the fractal dimension of the signal).

References

1. **Schur I.** On Power Series Which Are Bounded in the Interior of the Unit Circle // I. Schur. Methods in Operator Theory and Signal Processing, Ed. I. Gohberg, Operator Theory: Advances and Applications. – Vol. 18. – Birkhäuser-Verlag, 1986. – P. 31–60.
2. **Schur I.** On Power Series Which Are Bounded in the Interior of the Unit Circle II // I. Schur. Methods in Operator Theory and Signal Processing, Ed. I. Gohberg, Operator Theory: Advances and Applications. – Vol. 18. – Birkhäuser-Verlag, 1986. – P. 61–88.
3. **Lee D. T. L., Morf M., Friedlander B.** Recursive Least Squares Ladder Estimation Algorithm // IEEE Transactions on Circuits and Systems. – Vol. CAS-28. – No. 6, June 1981. – P. 78–86.
4. **Leon D. Iasemidis.** Epileptic Seizure Prediction and Control // IEEE Transactions on Circuits and Systems. – Vol. 50. – No. 5, May 2003. – P. 549–558.
5. **Liu H. S., Zhang T., Yang F. S.** A Multistage, Multimethod Approach for Automatic Detection and Classification of Epileptiform EEG // IEEE Transactions on Biomedical Engineering. – Vol. 49. – No. 12, December, 2002. – P. 1557–1566.
6. **Litt B., Echauz J.** Prediction of epileptic seizures // Lancet Neurology. – Vol.1. – 2001. – P. 22–30.
7. **Dingle A. A., Jones R. D., Carrol G. J., Fright W. R.** A Multistage System to Detect Epileptiform Activity in the EEG // IEEE Transactions on Biomedical Engineering. – Vol. 40. – No. 12, December, 1993. – P. 1260–1268.

Submitted for publication 2006 11 05

R. Ursulean, A. M. Lazar. Segmentation of Electroencephalographic Signals using an Optimal Orthogonal Linear Prediction Algorithm // Electronics and Electrical Engineering. – Kaunas: Technologija, 2007. – No. 1(72). – P. 73–76.

Finding the stationary parts of an EEG signal proves to be a difficult task because of the various shapes of the signal that can be found for different subjects, both normal and abnormal. The paper presents the results obtained in the segmentation of an EEG signal into stationary parts in order to be used in the stages of feature extraction and classification of these signals or in predicting epileptic seizures. For this purpose an optimal orthogonal linear prediction algorithm that implements a so-called "innovations filter" is used. The algorithm was tested on EEG signals from epileptic patients and the results are presented in detail. Some further applications and possible research fields are also offered. Ill. 8, bibl. 7 (in English; summaries in English, Russian and Lithuanian).

Р. Урсулеан, А. М. Лазар. Сегментация электроенцефалографических сигналов используя оптимальный ортогональный алгоритм линейного предсказания // Электроника и электротехника. – Каунас: Технология, 2007. – № 1(73). – С. 73–76.

Обнаружение постоянных частей сигнала электроэнцефалограммы достаточно трудная задача из-за различных форм сигнала, которые характерны для различных субъектов – и нормальных, и анормальных. Представлены результаты, полученные во время сегментации сигнала электроэнцефалограммы в постоянные части, чтобы использовать их в стадиях извлечения особенностей и классификации этих сигналов или в предсказании эпилептических припадков. С этой целью используется оптимальный ортогональный линейный алгоритм предсказания, который осуществляет так называемый "фильтр новшеств". Алгоритм был проверен на сигналах электроэнцефалограммы эпилептических пациентов. Подробно представлены результаты. Также предлагаются некоторые сферы исользования и возможные области дальнейшего исследования.Ил. 8, библ. 7 (на английском языке; рефераты на английском, русском и литовском яз.).

R. Ursulean, A. M. Lazar. Elektroencefalografinių signalų segmentavimas naudojant otimalų ortogonalų tiesinio prognozavimo algoritmą // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2007. – Nr. 1(73). – P. 73–76.

Stacionarias elektroencefalogramos dalis aptikti gana sunku, nes reikia tirti skirtingų formų signalus, gaunamus tiek iš normalių, tiek iš nenormalių subjektų. Pateikiami elektroencefalogramos segmentavimo į stacionarias dalis rezultatai, kurie bus naudojami tolesniuose etapuose, nustatant signalų savybes ir juos klasifikuojant, arba prognozuojant epilepsijos priepuolius. Šiam tikslui panaudotas optimalus ortogonalus tiesinio prognozavimo algoritmas, kuriame pritaikytas vadinamasis „inovacijų filtras". Algoritmo testavimas atliktas tiriant epilepsija sergančių pacientų elektroencefalogramas. Pateikiami išsamūs tyrimų rezultatai. Taip pat pasiūlytos kitos galimos taikymo sritys bei tolesnių tyrimų kryptys. Il. 8, bibl. 7 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).