

An Improved Signal Segmentation Method using Genetic Algorithm

Hamed Azami
M.Sc. Student
Department of Electrical
Engineering, Iran University
of Science and Technology,
Tehran, Iran

Karim Mohammadi
Full Professor
Department of Electrical
Engineering, Iran University
of Science and Technology,
Tehran, Iran

Hamid Hassanpour
Associate Professor
Department of School of
Information Technology and
Computer Engineering,
Shahrood University

ABSTRACT

In many signal processing application, the signal of interest is often divided into epochs. In these applications, the segmented signal is preferred to have no change on the statistical characteristics of the epochs. Modified Varri is among the segmentation methods with an acceptable accuracy. There are three parameters affecting on the accuracy of this method. These parameters are set experimentally. Hence, they may not be optimal for any signal segmentation application. We have used Genetic Algorithm (GA) in this research to choose appropriate values for these parameters in any signal segmentation application. The proposed technique was applied on both synthetic signal and Electroencephalography (EEG) to evaluate its performance. The results indicate superiority of the proposed method in signal segmentation compared to the original approach.

General Terms

Signal segmentation, Non-stationary Signal, Adaptive Segmentation, Genetic Algorithm (GA) and EEG Signal.

Keywords

Non-stationary Signal, Adaptive Segmentation, Modified Varri and Genetic Algorithm (GA).

1. INTRODUCTION

In a general manner signals are divided into two basic types: determined signal and random or statistical signal. Statistical signal is created by recorded random data, such as EEG and ECG signals. Depending on the behavior of signals, statistical signal is divided into stationary signal, which its statistical properties don't change over time, and non-stationary signal.

Because processing a stationary signal is easier than processing a non-stationary signal, in many signals processing, such as EEG analysis, a non-stationary signal is divided to the epochs which each epoch is a stationary signal. This act is called segmentation [1, 2].

Signal segmentation could be performed in two different ways, namely constant segmentation and adaptive segmentation. In constant segmentation a signal is segmented into epochs with constant duration [3, 4]. Although this method has a simple processing but its performance is extremely low. In the adaptive segmentation, signal boundaries are detected automatically. These boundaries depend on statistical characteristics such as

amplitude and frequency. There are a number of segmentation methods in the literature for adaptive segmentation [5-7].

In [8] it is suggested to use two scrolling windows to detect signal boundaries. These windows move over the entire signal and then in each window the data is modeled by using auto regressive model (AR). When both windows are in the same segment, the ARs for the two windows are similar. When the second window starts to move out of the segment, the AR difference between the two windows starts to increase. When the second window is completely out of the first segment, the difference between the AR of the two windows reach to maximum and the boundary is detected. This method is called GLR. To improve the generalized likelihood ratio (GLR) method another method used wavelet which is known as wavelet generalized likelihood ratio (WGLR) method [9].

Another technique for signal segmentation is Modified Varri technique. This method includes three parameters that must be determined experimentally. If we don't determine these parameters precisely, the boundaries of the segments won't be assigned correctly. In this paper to reach appropriate parameters, genetic algorithm (GA) which is a powerful search tool to look for exact or approximate solutions, is applied.

This paper is organized as follows. Next section explains Modified Varri and GA. In the third section the proposed method is described briefly. The performance of the suggested method is evaluated by synthetic data and real EEG signal which is represented in Section 4. Finally, Section 5 provides conclusion.

2. BACKGROUND KNOWLEDGE FOR THE PROPOSED METHOD

2.1 Modified Varri

There are two windows that move alongside the signal. These windows calculate two parameters: 1) amplitude measure (A_{dif}); and 2) frequency measure (F_{dif}); that are defined as below:

$$A_{dif} = \sum_{k=1}^l |x_k| \quad (1)$$

$$F_{dif} = \sum_{k=1}^l |x_k - x_{k-1}| \quad (2)$$

Where l and x_k show the window length and the k^{th} data point, respectively. The reason the two coefficients (A_l and F_l) are added to the equation is because one of the F_{dif} and A_{dif} always have greater effect on the G_m , so in order to balance their effect on G_m , A_l and F_l are used as follows [10]:

$$G_m = A_l |A_{dif_{m+1}} - A_{dif_m}| + F_l |F_{dif_{m+1}} - F_{dif_m}| \quad (3)$$

Where m represents number of the window. Local maximums in the G function, above a threshold that is defined before, indicate boundaries of each segment. It should be noted that adding A_l and F_l is proposed by Krajca [11].

2.2 Genetic Algorithm

The GA is a search technique based on natural selection and genetics that was proposed by Holland in 1975. GA is applied in a variety of significant problems, such as customizable FPGA IP core, the traveling salesman problem, network routing problems, stereo matching and clustering data [12-15]. The GA has four steps as follows:

Step 1: Initialization and calculation. First step in GA is generating a random population of strings that are called chromosomes, which encodes candidate solutions that are called individuals. To optimize a problem, GA develops toward better solutions and then computes fitness value of every member. Fitness function shows a value that depends on the distance between reached value and desired value.

Step 2: Selection. During each iteration, some of the chromosomes with higher fitness value is eliminated and then exchanged with new members. Fittest members that have the lowest fitness value which are called father and mother breed new members that are called children. Figure 1 shows how two chromosomes are recreated by parents [16]:

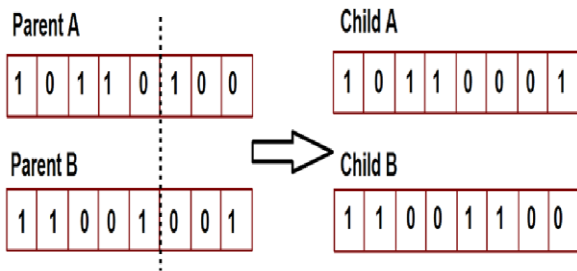


Fig 1: Children reproduction

Step 3: Mutation. Mutation can help the algorithm search which not to place in the local optimums. It can be applied by randomly flipping bits (or attribute values) within a single member. First a random number is generated for each bit of chromosome, if the random number is greater than the “mutation threshold” that is selected before this bit of chromosome is flipped.

Step 4: Termination. This process is repeated until a conclusion condition has been reached. There are many conditions for ending this algorithm such as when a solution is found that

satisfies optimum value or pre-defined number of iterations reached. Finally the flowchart of GA is presented in Figure 2 [6].

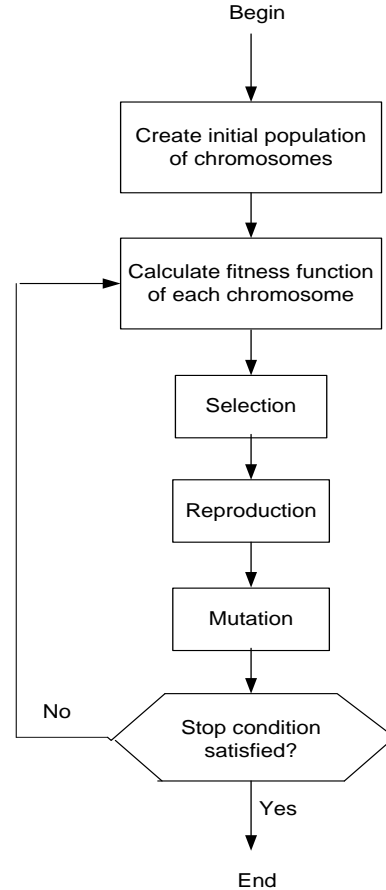


Fig 2: The flowchart of GA optimization cycle.

3. PROPOSED METHOD

First two windows move along the signal and then these windows calculate G_m with initial coefficients including window length, A_l and F_l . In this paper GA is used for attaining the best values for these variables. For this problem we defined fitness function as follows:

$$E_G = \frac{\sum_{m=0}^k |ceil(G_m - mean(G_m))|^2}{N} \quad (4)$$

Where N shows number of samples and k is the total number of analyzed windows. These parameters are used for GA: iteration=50, population=30. As described before, Local maximums in the G function, above a threshold that is defined before, indicate boundaries of each segment.

4. PERFORMANCE EVALUATION

WGLR and suggested method were implemented using Matlab R2009a from Math Works. The performance and efficiency of this method were evaluated using synthetic multi-component data, and real EEG data.

4.1 Synthetic Signal

For evaluating the performance of the suggested method, this algorithm is applied on a synthetic multi-component signal which each epoch was selected as a stationary signal as follows:

Epoch 1: $3.5\cos(2\pi t) + 4.5\cos(6\pi t)$,
 Epoch 2: $3.5\cos(2\pi t) + 5.5\cos(10\pi t)$,
 Epoch 3: $4.5\cos(2\pi t) + 5.5\cos(8\pi t)$,

Epoch 4: $4\cos(2.5\pi t) + 5.5\cos(5\pi t) + 2\cos(8\pi t)$,
 Epoch 5: $4\cos(4\pi t) + 6\cos(7\pi t)$,
 Epoch 6: $3\cos(3\pi t) + 7\cos(9\pi t)$,
 Epoch 7: $3\cos(3\pi t) + 2\cos(4\pi t) + 7\cos(8\pi t)$.

Figure 3.a and 3.b show 500 seconds of original signal and result of applying the proposed method, respectively. As can be seen the boundaries for all seven segments can be accurately detected. Figure 4.a and 4.b show 500 seconds of original signal and result of applying WGLR method, respectively. To demonstrate emphasis of this algorithm, in Figure 5, Modified Varri with arbitrary coefficients (window length=30, $A_I=7$ and $F_I=1$) is shown. The results show the proposed method has high performance for signal segmentation.

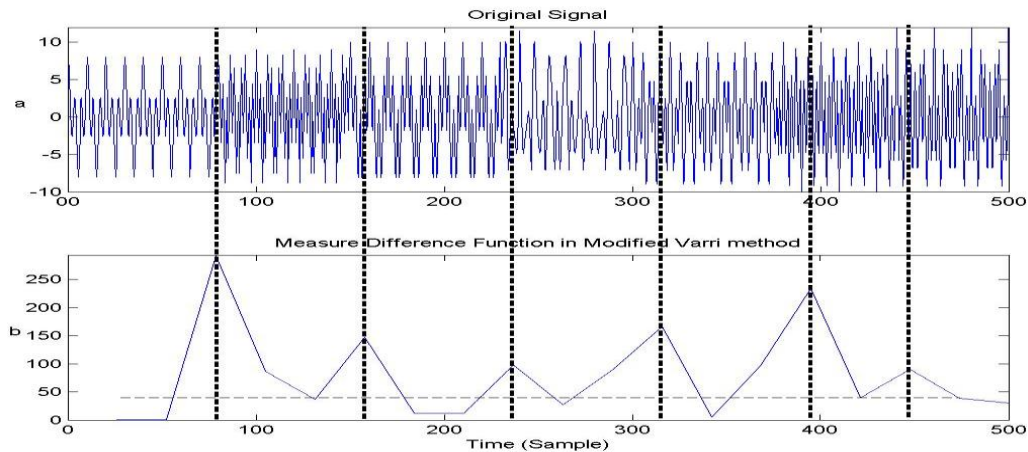


Fig 3: Signal segmentation in synthetic signal with using GA, (a) Original signal, (b) Output of the G function

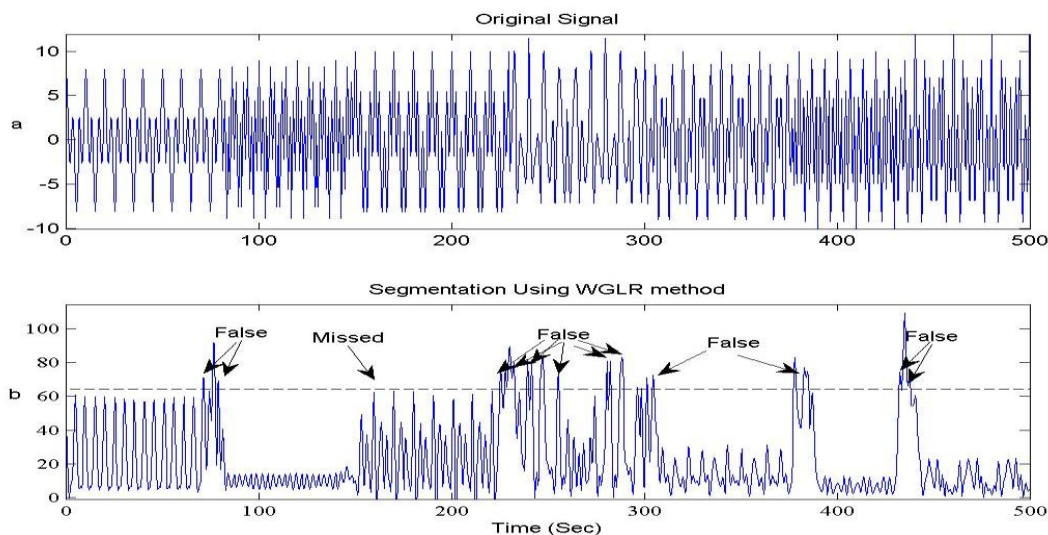


Fig 4: Signal segmentation in synthetic signal, (a) Original signal, (b) Output of the WGLR.

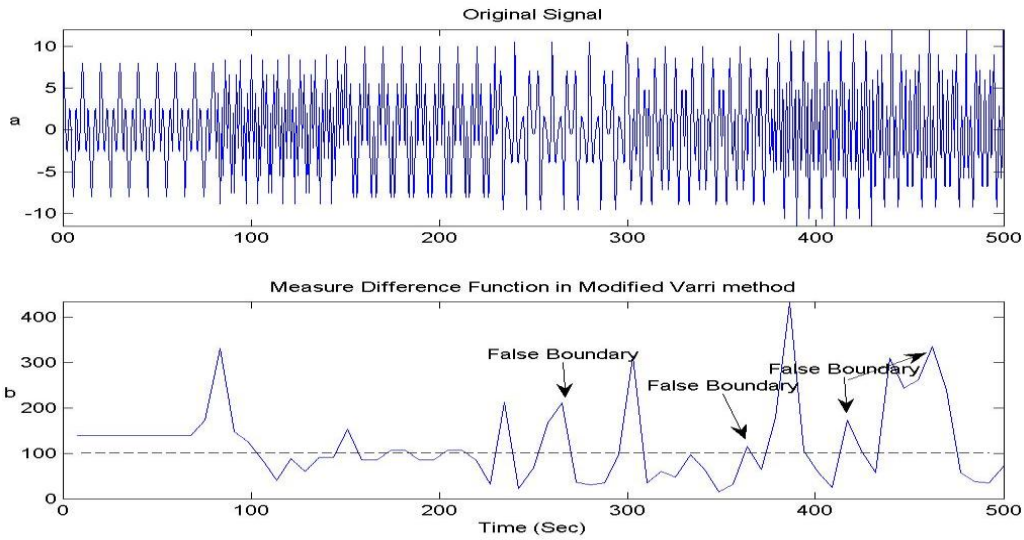


Fig 5: Signal segmentation in synthetic signal with arbitrary parameters, (a) Original signal, (b) Output of the G function

4.2. Real EEG

As described before, Signal segmentation is a pre-processing step for EEG signals. In this part we have used one epoch of a real newborn EEG signal that is shown in Figure 6.a. The result of applying Modified Varri with GA is shown in Figure 6.b. It can be seen that all seven segments can be accurately segmented.

In order to represent performance of this method, in Figure 7.a, WGLR is used for real newborn EEG signal just like it was used in Figure 6.a. Output of WGLR is shown in Figure 7.b. We can see the influence of this method on the achieved outputs.

5. CONCLUSION

One of the existing methods for signal segmentation is Modified Varri. This method has three parameters which must be defined experimentally. When these variables aren't defined correctly, the performance of this technique decreases drastically. In this paper a new technique for attaining parameters of Modified Varri method is introduced using genetic algorithm. The results indicate that the proposed method has better performance compared to the existing method.

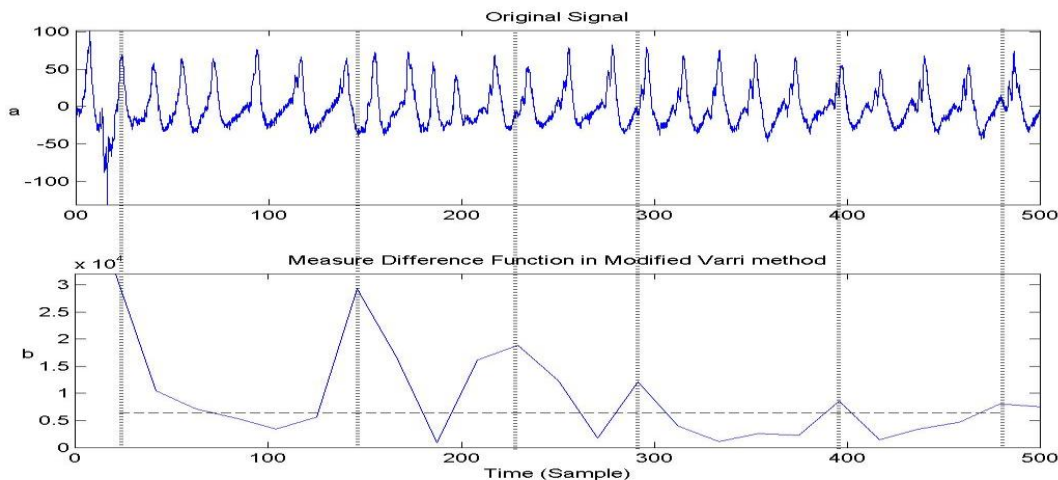


Fig 6: Signal segmentation in EEG signal with using GA, (a) Original signal, (b) Output of the G function.

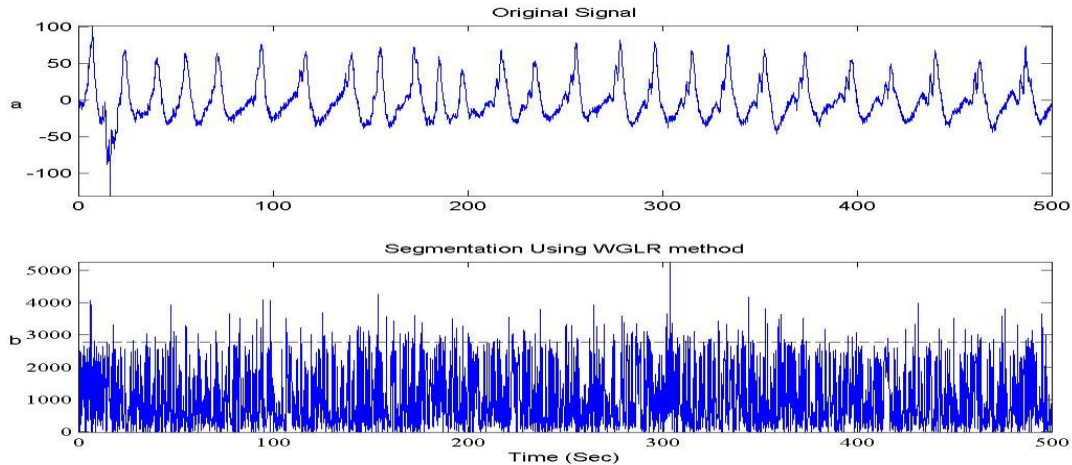


Fig 7: Signal segmentation in EEG signal, (a) Original signal, (b) Output of the WGLR.

6. REFERENCES

- [1] J. D. Scargle, “Studies in Astronomical Time Series Analysis. V. Bayesian Blocks, a New Method to Analyze Structure in Photon Counting Data”, *The astronomical journal*, vol. 504, pp. 405-418, 1998.
- [2] H. Hassanpour and M. Shahiri, “Adaptive Segmentation Using Wavelet Transform”, *International conference on Electrical Engineering*, pp. 1-5, April 2007.
- [3] K. Kosar, L. Lhotska, and V. Krajca, “Classification of Long-Term EEG Recordings”, *Lecture Notes in Computer Science*, vol. 3337, pp. 322-332, 2004.
- [4] R. Agarwal and J. Gotman, “Adaptive Segmentation of Electroencephalographic Data using a Nonlinear Energy Operator”, *Proc. IEEE International Symposium on Circuits and Systems (ISCAS'99)*, vol. 4, pp. 199-202, 1999.
- [5] M.E. Kirlangic, D. Perez, S. Kudryavtseva, G. Griessbach, G. Henning and G. Ivanova, “Fractal Dimension as a Feature for Adaptive Electroencephalogram Segmentation in Epilepsy”, *IEEE International EMBS Conference*, vol. 2, pp. 1573-1576, 2001.
- [6] S.M. Anisheh and H. Hassanpour, “Adaptive Segmentation with Optimal Window Length Scheme using Fractal Dimension and Wavelet Transform”, *International Journal of Engineering*, vol. 22, pp. 257-268, 2009.
- [7] H. Hassanpour and S.M. Anisheh, “An Improved Adaptive Signal Segmentation Method Using Fractal Dimension”, *IEEE Conference on Information Science, Signal Processing and their Applications*, pp. 720-723, 2010.
- [8] D. Wang, R. Vogt, M. Mason and S. Sridharan, “Automatic Audio Segmentation Using the Generalized Likelihood Ratio”, *2nd IEEE International Conference on Signal Processing and Communication Systems*, pp. 1-5, 2008.
- [9] J. Lv, X. Li and T. Li, “Web-based Application for Traffic Anomaly Detection Algorithm”, *Second IEEE International Conference on Internet and Web Applications and Services* pp. 44-60, 2007.
- [10] A. Varri, “Digital Processing of the EEG in Epilepsy”, *Licentiate Thesis, Tampere University of Technology, Tampere, Finland*, 1988.
- [11] V. Krajca, S. Petranek, I. Patakova and A. Varri, “Automatic Identification of Significant Graphoelements in Multichannel EEG Recordings by Adaptive Segmentation and Fuzzy Clustering”, *International Journal of Biomedical Engineering*, vol. 28, pp. 71-89, 1991.
- [12] Z. Zhang, C. Hou and J. Yang, “A Stereo Matching Algorithm based on Genetic Algorithm with Propagation Stratagem”, *IEEE Conference on Intelligent Systems and Applications*, pp. 1-4, 2009.
- [13] P. R. Fernando, S. Katkoori, D. Keymeulen, R. Zebulum, and A. Stoica, “Customizable FPGA IP Core Implementation of a General-Purpose Genetic Algorithm Engine”, *IEEE Transactions on Evolutionary Computation*, vol. 14, pp. 133-149, 2010.
- [14] M. Mohammadi, H. Alizadeh and B. Minaei-Bidgoli, “Neural Network Ensembles using Clustering Ensemble and Genetic Algorithm”, *IEEE Conference on Convergence and Hybrid Information Technology*, pp. 761-766, 2008.
- [15] J. J. Grefenstette, R. Gopal, B. J. Rosmaita, and D. V. Gucht, “Genetic algorithms for the Traveling Salesman Problem”, *International Conference of Genetic Algorithms*, pp. 160-168, 1985.
- [16] N. Ru and Y. Jianhua, “A GA and Particle Swarm Optimization Based Hybrid Algorithm”, *IEEE National basic research program of the China*, pp. 1047-1050, 2008.