Automatic Signal Segmentation Based on Singular Spectrum Analysis and Imperialist Competitive Algorithm

Hamed Azami, *Student Member, IEEE*
Department of Electrical Engineering
Iran University of Science and Technology
Tehran, Iran
hamed_azami@ieee.org

Saeid Sanei, *Senior Member, IEEE*
Faculty of Engineering and Physical Sciences
University of Surrey
Guildford, United Kingdom
s.sanei@surrey.ac.uk

***Abstract*— Electroencephalogram (EEG) is generally known as a non-stationary signal. Dividing a signal into the epochs within which the signals can be considered stationary, segmentation, is very important in many signal processing applications. Noise often influences the performance of an automatic signal segmentation system. In this article, a new approach for segmentation of the EEG signals based on singular spectrum analysis (SSA) and imperialist competitive algorithm (ICA) is proposed. As the first step, SSA is employed to reduce the effect of various noise sources. Then, fractal dimension (FD) of the signal is estimated and used as a feature extraction for automatic segmentation of the EEG. In order to select two acceptable parameters related to the FD, ICA that is a more powerful evolutionary algorithm than traditional ones is applied. By using synthetic and real EEG signals, the proposed method is compared with original approach (i.e. without using SSA and ICA). The simulation results show that the speed of SSA is much better than that of the discrete wavelet transform (DWT) which has been one of the most popular preprocessing filters for signal segmentation. Also, the simulation results indicate the performance superiority of the proposed method.**

Keywords- signal segmentation*; *singular spectrum analysis*; *fractal dimension*; *imperialist competitive algorithm

I. INTRODUCTION

Biomedical signals such as electroencephalogram (EEG) and electrocardiogram (ECG) are usually known as non-stationary, i.e. their statistical characteristics change over the time [1,2]. In such cases the signals are often segmented into smaller epochs during which the signal remains approximately stationary. The segmentation may be fixed or adaptive. Dividing non-stationary signals into fixed (rather small) size segments is easy and fast. However, it can not precisely follow the epoch boundaries [1-3]. On the other hand, in adaptive segmentation the boundaries are accurately and automatically followed [1]. Many adaptive segmentation methods have been suggested by researchers in the field such as those in [4-10]. Since noises can significantly decrease the performance of the segmentation methods, first we use singular spectrum analysis (SSA) as a filter. SSA is becoming an effective and powerful tool for time series analysis in meteorology, hydrology, geophysics, climatology, economics, biology, physics, medicine, and other sciences where short and long, one-dimensional and multi-dimensional, stationary and non-stationary, almost deterministic and noisy time series are to be analyzed [11]. The main advantage of SSA over principal component analysis (PCA) is that unlike PCA, SSA can cope with the cases where the number of signal components is more than the rank of the PCA covariance matrix. Then, fractal dimension (FD) is a measure of chaos which can indicate how large the changes in the signal statistics are. Katz's algorithm is a well-known technique to achieve a FD of a signal. Although the Hiaguchi's method computes the FD more precisely than Katz's method, due to sensitivity of Hiaguchi's method to noise, in signal segmentation, especially EEG signals, the Katz's method has been used [2]. Fig. 1 shows how FD changes when frequency and/or amplitude of a signal is changed.

The synthetic signal used here includes four segments. The first and second segments have the same amplitude. The frequency of the first part is however different from that of the second part. In the third segment the amplitude becomes different from that of the second segment. Amplitude and frequency in the 4th segment is different from those of the third segment. The reason for creating this signal is to show that if two adjacent epochs of a signal have different amplitudes and/or frequencies, the FD will vary.

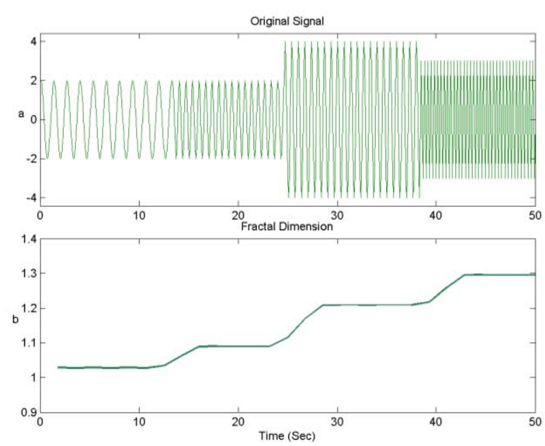

Figure 1. Variation of FD when amplitude or frequency changes.

Two effective parameters for FD that are determined empirically are called *window length* and *overlapping percentage* of the successive windows. Small windows may not be able to explain long-term statistics effectively and long windows ignore small block variations. The overlapping percentage of the successive windows affects both the computational load and the accuracy of the segmentation result.

In this paper, to achieve appropriate parameters, imperialist competitive algorithm (ICA) is employed. This algorithm is a quick search technique that can obtain exact or approximate local optimum responses in the desired space. ICA is a novel and powerful EA. From the results of their performances, ICA has been confirmed to perform better than the genetic algorithm (GA) [12].

The rest of the paper is organized as follows: In Section II.*A* SSA is explained. In Section II.*B* the Katz's method as a technique to calculate the FD is introduced. Section II.C introduces ICA briefly. The proposed adaptive method is explained in Section III. The performance of the proposed method is evaluated in Section IV. The last section concludes the paper.

II. Background Knowledge for the Proposed Method

A. *Singular Spectrum Analysis*

In this subsection a brief description of the two SSA stages together with the corresponding mathematics is given. At the first stage, the series is decomposed and at the second stage we reconstruct the original series and use the reconstructed series (which is without noise) to predict new data points [13].

1) Decomposition: This stage is composed of two sequential steps including embedding and singular value decomposition (SVD). In the embedding step, the time series **s** is mapped to k multidimensional lagged vectors of length l as follows:

$$\mathrm{x_i} = [s_{i-1}, s_i, \ldots, s_{i+l-2}]^T, \quad 1 \le i \le k \tag{1}$$

where $k = r - l + 1$, l is the window length ($1 \le l \le r$), and $[]^T$ denotes the transpose of a matrix. An appropriate window length totally depends on the application and the prior information about the signals of interest. The trajectory matrix of the series **s** is constructed by inserting each $\mathbf{x}_i$ as the ith column of an $l \times k$ matrix, i.e.

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k] = \begin{bmatrix} s_0 & s_1 & s_2 & \cdots & s_{k\text{-}1} \\ s_1 & s_2 & s_3 & \cdots & s_k \\ s_2 & s_3 & s_4 & \cdots & s_{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{l\text{-}1} & s_l & s_{l+1} & \cdots & s_{r\text{-}1} \end{bmatrix} \tag{2}$$

Note that the trajectory matrix **X** is a Hankel matrix, i.e. for all the elements along its diagonals i+j=constant.

In the SVD substage, the SVD of the trajectory matrix is computed and represented as the sum of rank-one biorthogonal elementary matrices. Consider the eigenvalues and corresponding eigenvectors of $\mathbf{S}=\mathbf{XX}^{\mathrm{T}}$ are $\lambda_1, \lambda_2, \ldots, \lambda_l$ and $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_l$, respectively. If $\mathbf{v}_i = \mathbf{X}^T \mathbf{e}_i / \sqrt{\lambda_i}$, then the SVD of the trajectory matrix can be written as

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \ldots + \mathbf{X}_d \tag{3}$$

where $d = \arg\max_i \{\lambda_i > 0\}$ and $\mathbf{X}_i = \sqrt{\lambda_i}\mathbf{e}_i \mathbf{v}_i^T$. The ith eigentriple of the SVD decomposition comprises of $\mathbf{v}_i$, $\mathbf{e}_i$, and λ_i. Projecting the time series onto the direction of each eigenvector yields the corresponding temporal principal component [14].

2) Reconstruction: This stage has two steps: grouping and diagonal averaging. The grouping step divides the set of indices {1, 2, ..., d} to m disjoint subsets I_1, I_2, ..., I_m. For every group $I_j = \{i_{j1}, i_{j2}, \ldots, i_{jp}\}$, we have $\mathbf{X}_{I_j} = \{X_{i_{j1}}, X_{i_{j2}}, \ldots, X_{i_{jp}}\}$. Grouping the eigentriples and expanding all matrices $\mathbf{X}_{I_j}$, (3) can be written as

$$\mathbf{X} = \{\mathbf{X}_{I_1}, \mathbf{X}_{I_2}, \ldots, \mathbf{X}_{I_m}\} \tag{4}$$

There is no general rule for grouping. For each application, the grouping rule depends on the special requirements of the problem and the type of the contributing signals and noise.

b) Diagonal averaging: In the final stage of analysis, each group is transformed into a series of length r. For a typical $l \times k$ matrix **Y**, the qth element of the resulted time series, g_q is calculated by averaging the matrix elements over the diagonal $i + j = q + 2$, where i and j are the row and column indices of **Y**, respectively [15].

The concept of separability is an important part of the SSA methodology. Assume that **s** is the sum of two series $\mathbf{s}_1$ and $\mathbf{s}_2$, i.e., $\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2$. Separability means that the matrix terms of the SVD of the trajectory matrix of X can be divided into two disjoint groups, such that the sums of the terms within the groups result in the trajectory matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ of the time series $\mathbf{s}_1$ and $\mathbf{s}_2$, respectively [14]. A necessary condition for separability of the sources is disjointness of their frequency spectrum. It is also worth mentioning that exact separability cannot be achieved for real-world signals; hence, only approximate separability can be considered.

The eigentriples resulting from the SSA also contain information about the frequency content of the data. If there is a periodic component in the data, it will be reflected in the output of the SSA as a pair of (almost) equal eigenvalues [13]. Moreover, the highest peaks in the Fourier transform of the corresponding eigenvectors are related to the frequency of the periodic component. These features of the SSA are used to construct data-driven filters [15].

B. Fractal Dimension

In Euclidean space, line and page are known as one dimensional and two dimensional, respectively and non-integer dimension does not exist, but FD represents a non-integer dimension regarding to the concepts of modern mathematics. It is commonly used in analysis of biomedical signals such as EEG and ECG, image processing and electrochemical processes [16-18]. There are some methods to calculate the FD of a signal such as Hiaguchi, Petrosian, and Katz's methods [19]. Because FD is directly estimated from the time-varying signal, it has low computational cost. Katz's algorithm has a lower sensitivity to noise and good speed in contrast to the two other algorithms [18]. Using Katz's algorithm:

$$FD = \frac{\log_{10}(n)}{\log_{10}(d/L) + \log_{10}(n)} \tag{1}$$

where L is sum of the distances between consecutive points and d is the maximum distance between the first data of time series and data that has maximum distance from it. Also, $n=L/a$ shows the step size in time series [20].

C. Imperialist Competitive Algorithm

ICA is a novel population based optimization algorithm proposed in 2007 by Atashpaz-Gargari and Lucas [12]. Today this algorithm has many applications such as designing controller for industrial systems, solving optimization problems in PID controller, communication systems, and training and analysis of artificial neural networks [21-24].

Like other evolutionary algorithms, this algorithm begins with the initial population with random numbers that each of them is called a "Country". Some of the members of the population that have best fitness values are selected as imperialists. Each member of the remaining population is called a colony. Total fitness values of an empire relies on both the power of the imperialist country and the power of its colonies.

In each stage, the countries move toward their related imperialist. If the fitness value of a colony achieves more than its related fitness value of imperialist then, this colony and its related imperialist transform to imperialist and colony, respectively. In every stage, the weakest colony of the weakest empire moves toward the closest empire, and the empire without any colony is eliminated. After a while, all empires fall down except for the most powerful one and all the colonies go under the control of this unique empire [25]. Finally the flowchart of the corresponding ICA is presented in Fig. 2.

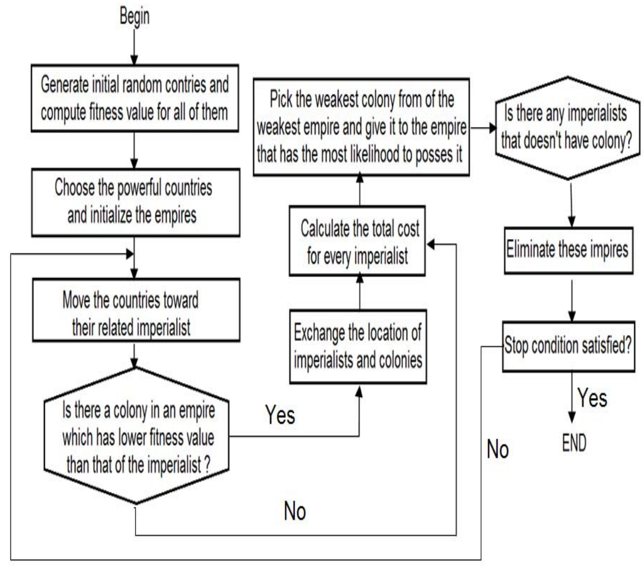

Figure 2. Flowchart of the ICA

III. PROPOSED ADAPTIVE SEGMENTATION

In this section the proposed method is explained briefly in four steps as follows:

1. The original signal is initially filtered using the SSA. The filtered signal can indicate the slowly changing features of the signal in the lower frequency bands. In addition, the SSA has no shifting effect on the patterns existing on the time-series after filtering compared with other FIR filters. Also, the speed of SSA is higher than discrete wavelet transform (DWT) that was used for signal segmentation.
2. Two successive windows are slide along the signal as is shown in Fig. 3. For each window, the FD is computed using the Katz's method. In order to detect the segments boundaries, the FD variations are computed as follows

$$G_t = \left|FD_{t+1} - FD_t\right|, \qquad t = 1,\ 2,\ ...,\ L\text{-}1 \tag{6}$$

where t and L are the number of analyzed windows and the total number of analyzed windows, respectively. If the windows are placed in a segment, their statistical properties don't change, in other words the achieved FDs remain approximately constant and equal. However, as can be seen in Fig. 3 when sliding windows fall in the different segments, the FD of them changes and the boundary is detected.

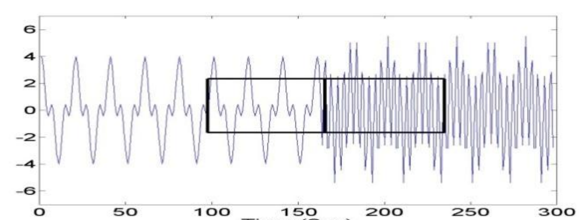

Figure 3. State of placing of the sliding windows

3. As described before, we use two parameters that affect on accuracy of the signal boundaries, namely, the length of the window and percentage overlapping of the successive windows and if these parameters are not chosen appropriately, the segments boundaries may be inaccurate. In this part we use ICA. Fitness function of the ICA over k shifts of the sliding window is chosen as:

$$E_G = \frac{\sum_{t=0}^{k} \left| ceil(G_t - mean(G)) \right|^2}{N} \qquad (7)$$

where N shows the number of samples in G and *ceil* stands for ceiling. It has been shown that using the above cost function leads to more accurate segment boundaries compared with the method introduced in [9,25]. As mentioned before, ICA can optimize a fitness function better than the previous optimization algorithms such as GA.

A deficiency of the method introduced in [26] is that percentage overlapping of the successive windows has been adjusted empirically. In our research, E_G has been defined as a function of window length, percentage overlapping of the successive windows, and mean value of G and these parameters can be selected automatically.

4. Mean value of G ($\overline{G}$) is chosen as the threshold level, meaning that when the local maximum is bigger than this threshold, the current time sample is chosen as a boundary of the segment.

IV. SIMULATION RESULTS

In order to evaluate the performance of the suggested method, we use two kind signals, namely, the synthetic data and real EEG signals. The synthetic signal includes the following seven epochs:

Epoch 1: 0.5cos(πt) + 1.5cos(4πt) + 4cos(5πt),
Epoch 2: 0.7cos(πt) + 2.1cos(4πt) + 5.6cos(5πt),
Epoch 3: 1.5cos(2πt) + 4cos(8πt),
Epoch 4: 1.5cos(πt) + 4cos(4πt),
Epoch 5: 0.5cos(πt) + 1.7cos(2πt) + 3.7cos(5πt),
Epoch 6: 2.3cos(3πt) + 7.8cos(8πt),
Epoch 7: 0.8cos(πt) + cos(3πt) + 3cos(5πt).

In Fig. 4.a and 4.b the synthetic data described above and the filtered signal by SSA with window length 9 are shown, respectively. As we can see, the filtered signal is smoother than the original signal. As mentioned before, two parameters are needed to calculate the FD of a filtered signal, window length and percentage overlapping of the successive windows, which affect the performance of the proposed approach. In this paper, these parameters are optimally set using the ICA. ICA which inspired from principles of natural evolution has been popular in last for decades in order to optimization problems in many engineering and natural science applications. ICA applied here has 45 populations and the number of iterations is 50. Also length of the window and the percentage of overlap for ICA are selected between 2% and 15% of the signal length with trial and error.

Fig. 4.b, Fig. 4.c and Fig. 4.d illustrate smoothed or filtered signal by SSA, the FD of the filtered signal and the changes in G function, respectively. In order to emphasize on using SSA as a preprocessing step in this method, in Fig. 5, the result of the simple fractal method without the SSA is shown. As can be seen in Fig. 5, the proposed method can not reveal one segment boundary accurately (miss boundary). Also, the method detects two boundaries incorrectly (false boundary).

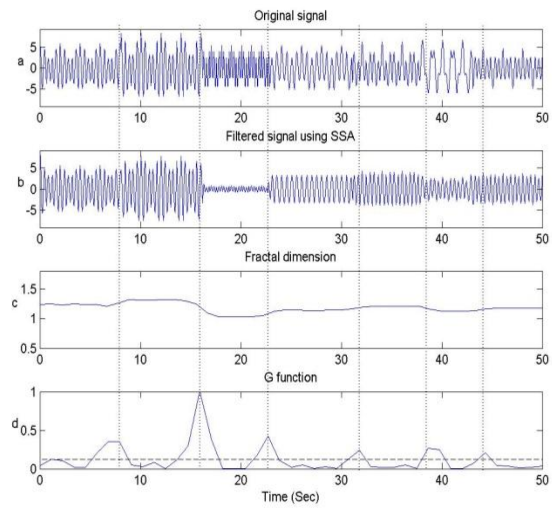

Figure 4. Result of the proposed technique; (a) original signal, (b) filtered signal by SSA, (c) output of FD, and (d) G function result. As can be seen the boundaries for all seven segments can be accurately detected.

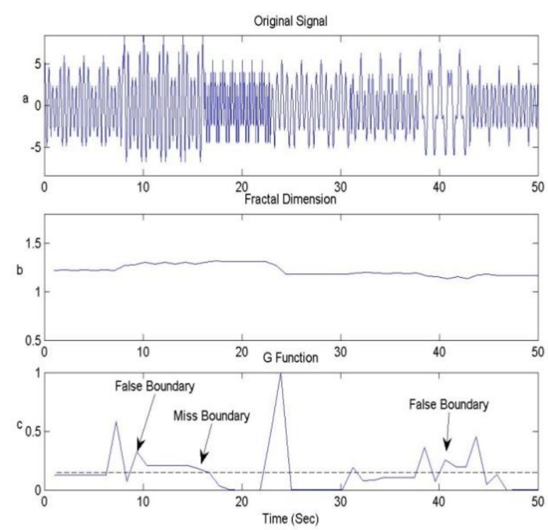

Figure 5. Signal segmentation in synthetic signal; (a) original signal, (b) fractal dimension of the original signal, and (c) output of G function.

EEG is an electrical activity used as an important tool by physicians to study the functional state of the brain and to diagnose certain neurophysiologic disorders [1]. As explained before, signal segmentation is a pre-processing step for EEG signals. Fig. 6.a shows a real newborn EEG signal which the length of this signal and the sampling frequency are 7500 samples and 256 Hz, respectively. The result of applying the proposed method is shown in Fig. 6.d. The SSA used for EEG signals of this paper has window length equal to 40. Also, we test our proposed method by using another EEG signal in Fig. 7. In Fig. 6.d and Fig. 7.d can be seen that all segments boundaries of the signals are accurately selected.

In order to emphasize on using the SSA as a preprocessing step, the results of the simple fractal method without using the SSA corresponding to the Fig. 6 and 7 are shown in Fig. 8 and 9, respectively. As can be seen in Figure 8.c/9.c, the proposed method without using SSA incorrectly detects one boundary of the segment (false boundary).

In this paper the simulations have been carried out using a DELL-PC with Intel (R) Core (TM) i3 CPU M350 2.27 GHz and 2-GB RAM by MATLAB R2010a. The CPU times for computing DWT and SSA for the EEG signal are gotten 0.3827 s and 0.0574 s, respectively. Therefore, using SSA is much better than using DWT that was used in traditional segmentation methods [26].

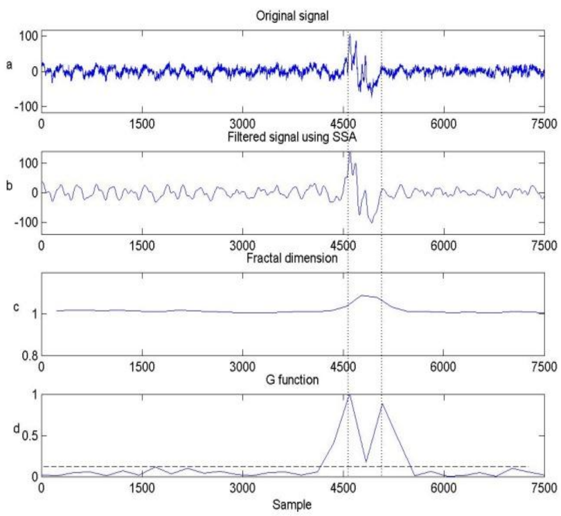

Figure 6. Signal segmentation in real EEG signal; (a) original signal, (b) filtered signal using SSA, (c) fractal dimension of the filtered signal, and (d) output of G function.

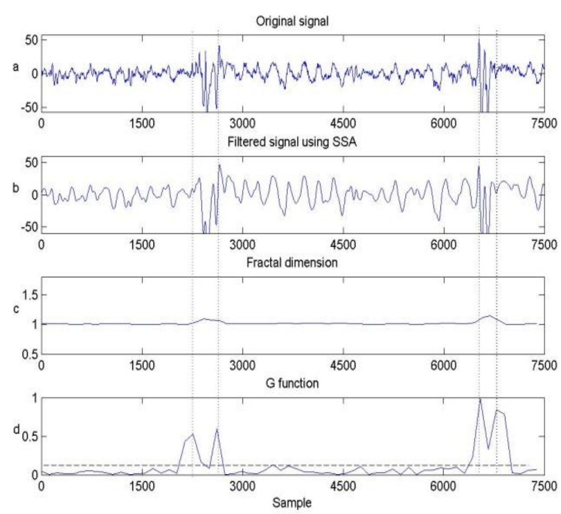

Figure 7. Signal segmentation in real EEG signal; (a) original signal, (b) filtered signal using SSA, and (c) fractal dimension of the filtered signal, and (d) output of G function.

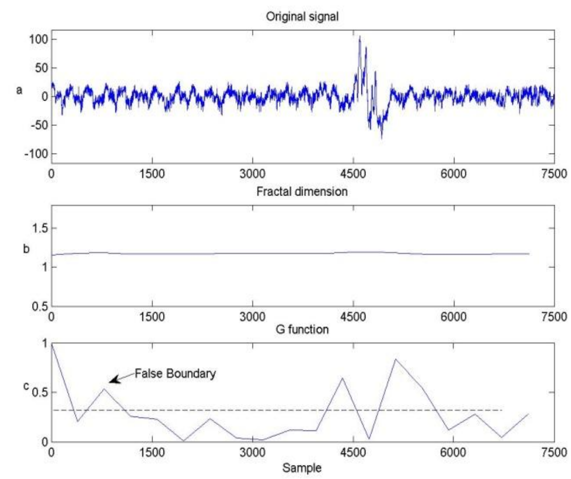

Figure 8. Signal segmentation in synthetic signal; (a) original signal, (b) fractal dimension of the original signal, and (c) output of G function.

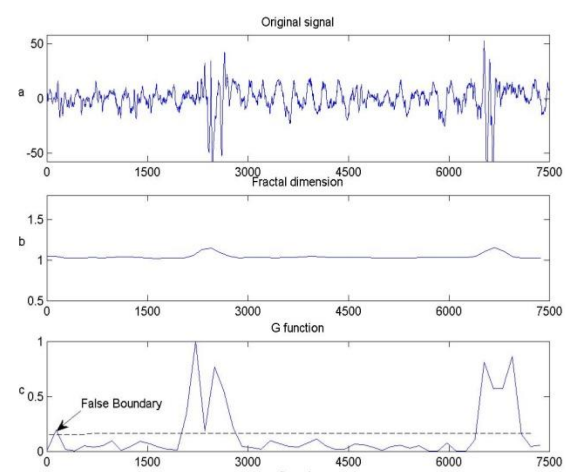

Figure 9. Signal segmentation in synthetic signal; (a) original signal, (b) fractal dimension of the original signal, and (c) output of G function.

V. CONCLUSIONS

In this paper an improved adaptive segmentation method using SSA, FD and ICA has been introduced. The new hybrid structure provides a stronger evolutionary technique for analysis of non-stationary data such as EEG signals. Since noises can significantly affect the performance of the segmentation methods, the SSA as a powerful tool for mitigation of noises has been used. Also, the CPU time for SSA is 6 times lower than that of the DWT as traditional method for signal segmentation. This plays an important role in signal segmentation, particularly in online applications. After filtering the signal, the FD by Katz's method has revealed the changes in both the amplitude and frequency of the signal. Finally, to optimize the selection of parameters of the FD, ICA as a powerful tool has been used. The results indicate superiority of the proposed method for segmenting the signals, especially EEGs.

REFERENCES

[1] H. Azami, S. Sanei, and K. Mohammadi, "A novel signal segmentation method based on standard deviation and variable threshold," Journal of Computer Applications, vol. 34, no. 2, pp. 27-34, 2011.

[2] H. Azami, B. Bozorgtabar, and M. Shiroie, "Automatic signal segmentation using the fractal dimension and weighted moving average filter," Journal of Electrical & Computer science, vol. 11, no. 6, pp. 8-15, 2011.

[3] R. Agarwal and J. Gotman, "Adaptive segmentation of electroencephalographic data using a nonlinear energy operator," IEEE International Symposium on Circuits and Systems (ISCAS'99), vol. 4, pp. 199-202, 1999.

[4] H. Hassanpour, M. Mesbah, and B. Boashash, "Time-frequency based newborn EEG seizure detection using low and high frequency signatures," Physiological Measurement. vol. 25, pp. 935-944, 2004.

[5] H. Hassanpour, M. Mesbah, and B. Boashash, "Time-frequency feature extraction of newborn EEG seizure using SVD-based techniques," EURASIP Journal on Applied Signal Processing, vol. 16, pp. 2544-2554, 2004.

[6] K. Kosar, L. Lhotska, and V. Krajca, "Classification of long-term EEG recordings," Lecture Notes in Computer Science, vol. 3337, pp. 322-332, 2004.

[7] 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.

[8] H. Azami, K. Mohammadi, and B. Bozorgtabar, "An improved signal segmentation using moving average and Savitzky-Golay filter," Journal of Signal and Information Processing, vol. 3, no. 1, pp. 39-44, 2012.

[9] H. Azami, K. Mohammadi, and H. Hassanpour, "An improved signal segmentation method using genetic algorithm," Journal of Computer Applications, vol. 29, no. 8, pp. 5-9, 2011.

[10] H. Hassanpour and M. Shahiri, "Adaptive segmentation using wavelet transform," International Conference on Electrical Engineering, pp. 1-5, April 2007.

[11] S. Sanei, T. K. M. Lee, and V. Abolghasemi, "A new adaptive line enhancer based on singular spectrum analysis," IEEE Transaction on Biomedical Engineering, vol. 59, no. 2, 2012.

[12] E. Atashpaz-Gargari and C. Lucas, "Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition," IEEE Congress on Evolutionary Computation, pp. 4661-4666, 2007.

[13] H. Hassani, "Singular spectrum analysis: methodology and comparison," Journal of Data Science, vol. 5, no. 2, pp. 239-257, 2007.

[14] F. Ghaderi, H. Mohseni, and S. Sanei, "Localizing heart sounds in respiratory signals using singular spectrum analysis," IEEE Transactions on Biomedical Engineering, vol. 59, no. 12, pp. 3360-3367, 2011.

[15] Y. Tao, E. C.M. Lam, and Y. Y. Tang, "Feature extraction using wavelet and fractal," Pattern Recognition, vol. 22, no. 3-4, pp. 271-287, 2001.

[16] U. R. Acharya, O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan, "Non-linear analysis of EEG signals at various sleep stages," Computer Methods and Programs in Biomedicine, vol. 80, no. 1, pp. 37-45, 2005.

[17] W. Y. Hsu, C. C. Lin, M. S. Ju and Y. N. Sun, "Wavelet-based fractal features with active segment selection: application to single-trial EEG data," Neuroscience Methods, vol. 163, no. 1, pp. 145-160, 2007.

[18] R. Esteller, G. Vachtsevanos, J. Echauz, and B. litt, "A comparison of fractal dimension algorithms using synthetic and experimental data," IEEE International Symposium on Circuits and Systems (ISCAS'99), vol. 3, PP. 199-202, 1999.

[19] R. Esteller, G. Vachtsevanos, J. Echauz, and B. litt, "A comparison of waveform fractal dimension algorithms," IEEE Transaction on Circuits and Systems, vol. 48, no. 2, pp. 177-183, 2001.

[20] M. J. Katz, "Fractals and the analysis of waveforms," Computers in Biology and Medicine, vol. 18, no. 3, pp. 145-156, 1988.

[21] M. Abdechiri, K. Faez, and H. Bahrami, "Neural network learning based on chaotic imperialist competitive algorithm," 2nd International Intelligent Systems and Applications (ISA), pp. 1-5, 2010.

[22] E. Atashpaz-Gargari, F. Hashemzadeh, R. Rajabioun and C. Lucas, "Colonial competitive algorithm: a novel approach for PID controller design in MIMO distillation column process," International Journal of Intelligent Computing and Cybernetics, Vol. 1, No. 3, pp. 337-355, 2008.

[23] R. Rajabioun, E. Atashpaz-Gargari, and C. Lucas, "Colonial competitive algorithm as a tool for Nash equilibrium point achievement," Book Chapter, Springer Lecture Notes in Computer Science, pp. 680–695, 2008.

[24] E. Atashpaz-Gargari, R. Rajabioun, F. Hashemzadeh, and F. Salmasi, "A decentralized PID controller based on optimal shrinkage of gershgorin bands and PID tuning using colonial competitive algorithm," International Journal of Innovative Computing, Information and Control, vol. 5, no. 10, pp. 3227-3240, 2009.

[25] S. M. Anisheh, "An adaptive segmentation method using fractal dimension and wavelet transform," Journal of Computer Engineering, vol. 1, pp. 11-18, 2009.

[26] 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.