Singular Spectrum Analysis Based EMG Artifact Removal from ECG Signal

Animesh Sarangi[1], Bal Gopal Mishra[2], Satyabhama Dash[3]
[1]*Post Graduate Scholar, EIE Department, Odisha University of Technology and Research, Odisha, India.*
[2]*Post Graduate Scholar, EIE Department, Odisha University of Technology and Research, Odisha, India.*
**Corresponding author: sarangianimesh1997@gmail.com*
[3] *Assistant Professor, EIE Department, Odisha University of Technology and Research, Odisha, India.*

Abstract

Electromyogram (EMG) or muscle artifacts frequently affect electrocardiogram (ECG) readings. These artifacts make the required information in the ECG signal difficult to see. In this study, we introduced the singular spectrum analysis (SSA), a powerful subspace-based method for removing EMG artifacts from ECG data. In order to effectively extract the desired component from the tainted ECG data, we presented a new grouping approach and set a threshold. First, a process known as embedding converts a single channel signal into several channels of signals or data. The orthogonal eigenvectors are then calculated using singular value decomposition(SVD) from the multichannel data's covariance matrix. A threshold is selected to locate these eigenvectors, which are utilized to generate the required subspace. After locating the subspace, the multichannel data is simply projected into it, followed by a method called diagonal averaging which will create the original time series and extract the ECG signals.

Keywords: *Electrocardiogram, EMG artifact, Singular Spectrum Analysis, Embedding, SVD, Mobility.*

1. Introduction

The ECG is a discreet demonstration tool used to capture the electrical activity of the heart. This is done by measuring the difference between the electrodes that are placed on the skin at certain locations on the human body. Every heartbeat causes the chamber and ventricle to depolarize and repolarize, according to one Electrocardiographic (ECG) pattern. According to Fig. 1, ECG signals consist of P, QRS, and T segments that are spaced apart by isoelectric zones.

The main issue with ECG recording is the disturbances or artifacts that can result from baseline wandering, Electromyogram (EMG) artifact, powerline interference,loose electrode artifact and motion artifacts, which can hamper the signal quality and lead to incorrect diagnosis [1] [2]. When a subject's ECG is being recorded while they are going about their regular business,the movement of the skin under the electrodes causes motion artifacts in the ECG. The P, QRS, and T segments of the ECG signal have frequency components that coincide with the spectrum of motion artifacts [3]. However, EMG interference results in certain high frequency motion distortions. In the ECG signal these distortions might alter the cardiac characteristics' amplitudes or manifest as extra segments. The frequency spectrum overlap between the ECG signal and these motion distortions makes conventional filtering techniques ineffective [4].

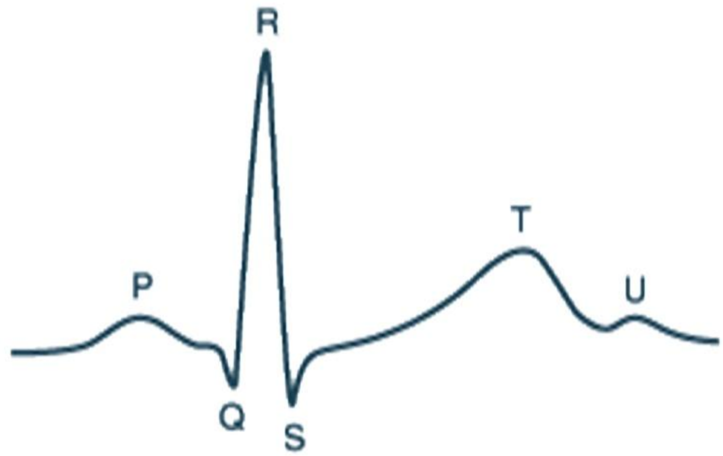

Figure 1. ECG signal [5]

This work describes a quick and effective method for removing muscle or EMG abnormalities from the ECG data known as singular spectrum analysis (SSA). The single channel signal in SSA is first projected into a data matrix (X), a process known as embedding. After that, the orthogonal eigenvectors are evaluated by singular value decomposition (SVD). We initially assessed each eigenvector's first-order local variations in order to extract the relevant component. The data matrix X is then projected into the space encompassed by the eigenvectors whose local fluctuations are fewer than the already set threshold by establishing the appropriate threshold. The artifact-free ECG signal is then recovered by running the data matrix X through the reverse embedding procedure and diagonal averaging. The suggested approach is evaluated using real-world ictal ECG data affected by EMG artifacts and synthetic sinusoidal signals contaminated by random noise. Simulation findings demonstrate that the suggested strategy effectively gets rid of EMG artifacts.

2. Methodology

The aim of SSA is to break down the original series into a manageable number of isolated, comprehensible parts. Decomposition and reconstruction are the two complementing processes that make up the SSA approach. Each stage consists of two distinct procedures. We break down the series in the first step, and then we rebuild the original series in the second. The graphical representation of all the methods involved in SSA is given in Fig. 2.

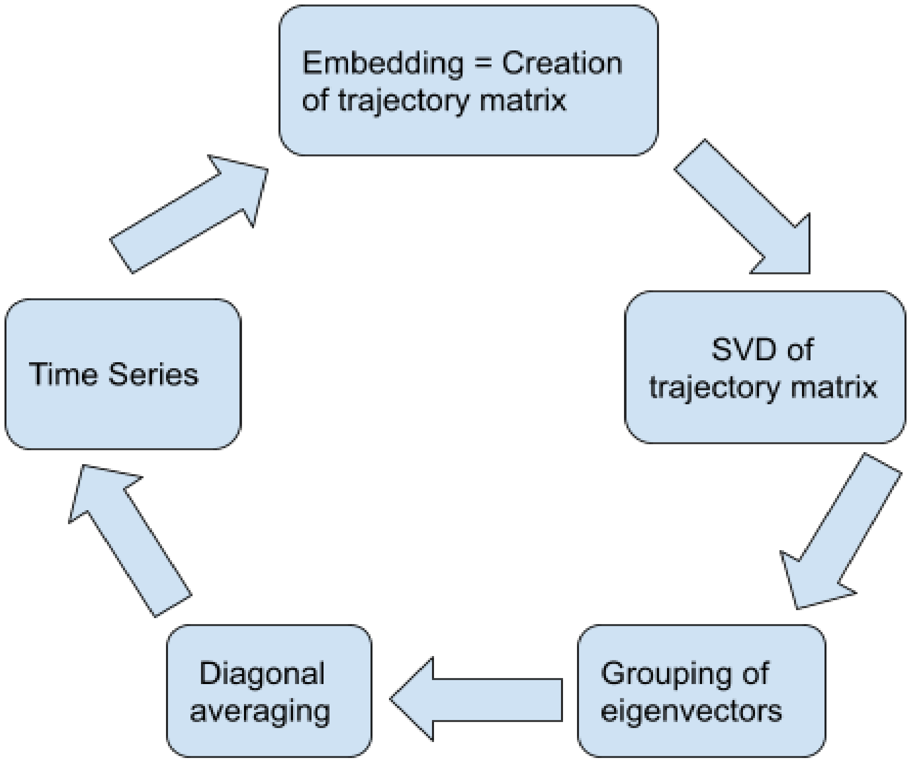

Figure 2. Graphical representation of steps involved in SSA

2.1. Decomposition

2.1.1 Embedding: A one-dimensional time series, $Y_T = (y_1, y_2, y_3,...., y_T)$, can be mapped into a multidimensional series, $X_1, X_2, X_3,...., X_K$, using a vector, $X_i = (y_i, y_{i+1}, y_{i+2},...., y_{i+L-1})^T \in R^L$, where L = T - K +1. L-lagged vectors are referred to as X_i vectors. A suitable window length L should be chosen so that $2 \leq L \leq T$, serves as the embedding only input. $X = [X_1, X_2, X_3,...., X_K]^T = (x_{ij})_{i,j=1}^{L,K}$ is the end result of this phase which is known as the trajectory matrix . This trajectory matrix is referred to as a Hankel matrix since each ascending skew diagonal is constant from left to right i.e i+j = constant.

2.1.2 Singular Value Decomposition: In order to get the orthogonal eigenvectors,first we have to factorizes the trajectory matrix into three matrices, $X = USV^T$, where S is a diagonal matrix,and U and V are orthogonal matrices. The appropriate eigenvalues and eigenvectors are represented as $\lambda_1, \lambda_2, \lambda_3,...., \lambda_L$ and $u_1, u_2, u_3,...., u_L$ in the covariance matrix D for the data matrix X provided by $D = XX^T$, respectively. Here the eigenvalues and accompanying eigenvectors are listed in descending order.

Here V_i can be found by

$$V_i = \frac{X^T U_i}{\sqrt{\lambda_i}} \quad (1)$$

where i = 1,2, .. , L. So the SVD of trajectory matrix X can be written as

$$X = X_1 + X_2 + \dots + X_L = \sum_{i=1}^{L} X_i \quad (2)$$

where, $X_i = \sqrt{\lambda_i} U_i V_i^T$

2.2 Reconstruction

2.2.1 Grouping: The biggest challenge in SSA is that in order to get the required ECG signal appropriate subspace must be chosen for which eigenvalues should be selected manually. There are various grouping methods available to select the desired eigenvalues and corresponding eigenvectors. Here we have used the mobility or local variation method.This method differs from previous ones in that the desired eigenvectors are chosen according to their local fluctuations rather than the magnitude of their corresponding eigenvalues [6].

2.2.2 Mobility or Local fluctuation method: Consider a 'T' length signal vector $z = [z_1, z_2, z_3, \dots, z_K]$. The local fluctuation of the signal m_Z [7] is found by,

$$m_z = \frac{\sqrt{\frac{\sum_{j=1}^{T} z(j)^2}{T}}}{\sqrt{\frac{\sum_{j=1}^{T-1} d(j)^2}{T-1}}} \quad (3)$$

where d(j) = z(j) - z(j - 1) is the first order variation of z(j) signal. The mobility m_z for the EMG and ECG signals are high and low respectively, since the EMG signal seems to be random.

So the eigenvectors having local mobilities below the predetermined threshold span the subspace onto which the X matrix is projected.

2.2.3 Diagonal averaging: In the final step diagonal averaging method is performed on the trajectory matrix which will convert the matrix X into a time series of length T i.e we will get the required EMG artifact free ECG signal.

3. Results

3.1 Data Collection and Description

The Beth Israel Hospital Arrhythmia Laboratory acquired approximately 4000 long-term Holter recordings, which are the ECGs' source featured in the MIT-BIH Arrhythmia Database[8]. These recordings came from inpatients in around 60% of the cases. 23 entries, which are picked at random from this batch and are marked from 100 to 124 inclusively (with some numbers missing), are included in the database. These records all last a little over 30 minutes. In our work we have taken the ECG sample of patient number 103 who is a male having a heartbeat in the range of 62-92 Beats Per Minute(BPM).

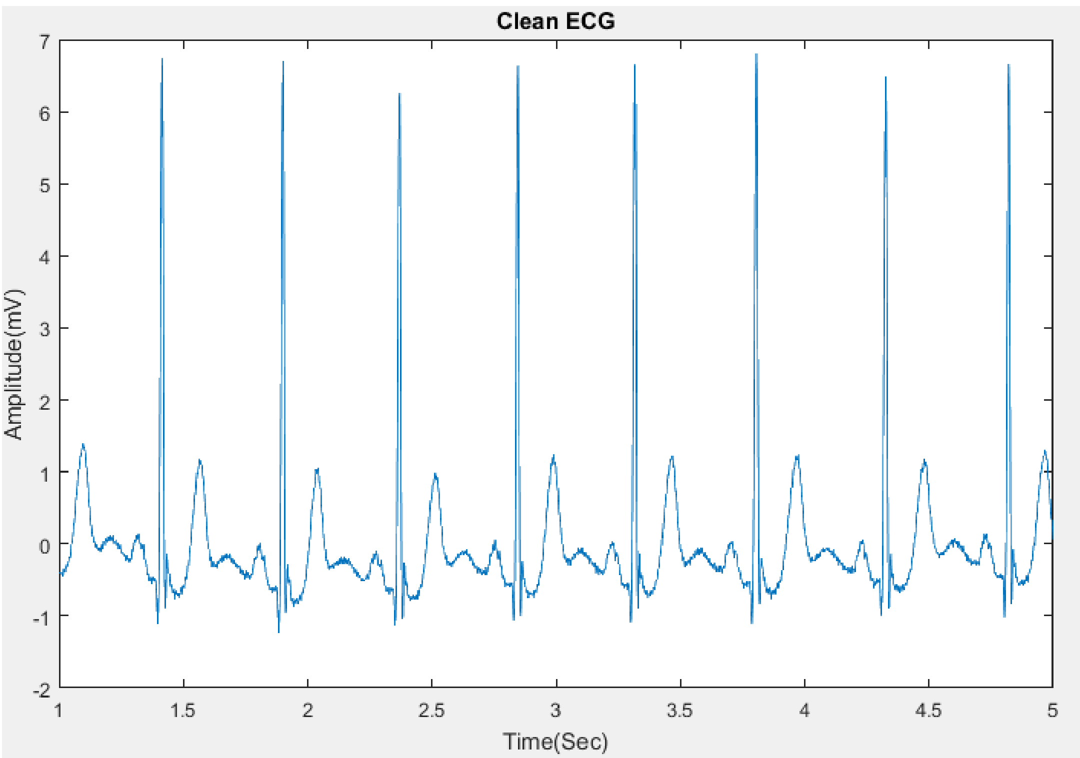

Figure 3. Patient's noise free ECG signal

3.2 ECG signal extraction by removing EMG artifacts

We treat an ECG signal with a length of 5 sec as X(t) having frequency 512 Hz which can be seen in Fig.3. The noisy signal Ns(t) = X(t)+N(t) is created by adding the noise N(t), which represents the EMG signal or artifact. Fig.4. shows the effect of EMG artifacts on the ECG signal.

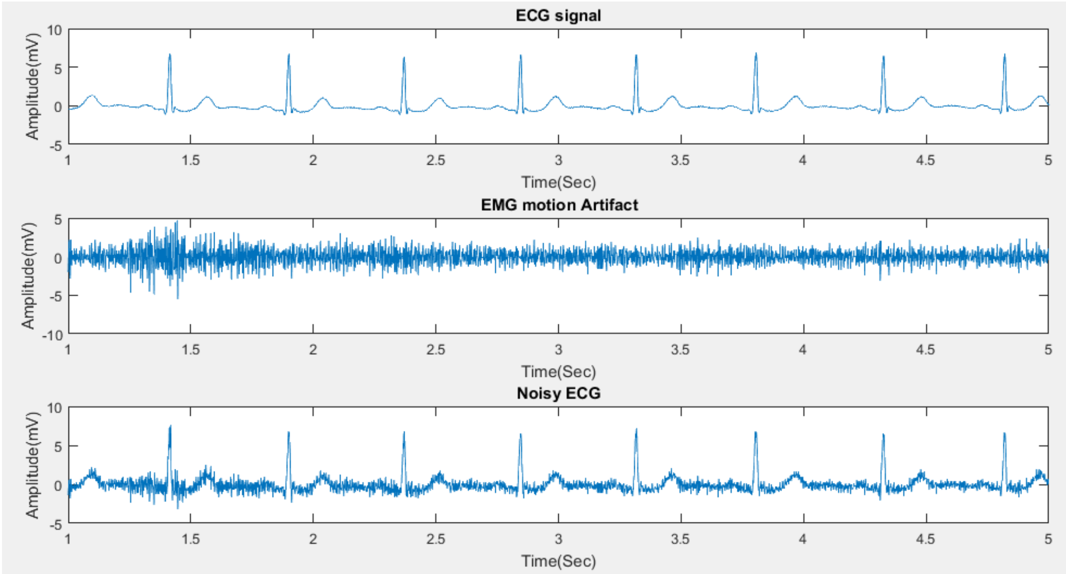

Figure 4. Addition of EMG artifacts in ECG signal

For all simulations, the window length L—which is necessary to transform a single-dimensional signal data into a multidimensional signal—is slated to 13; this value is chosen using the criteria outlined in [6]. After the mobility of the eigenvectors is calculated, all the mobility of the respective eigenvectors are shown in a graphical manner in Fig.5.

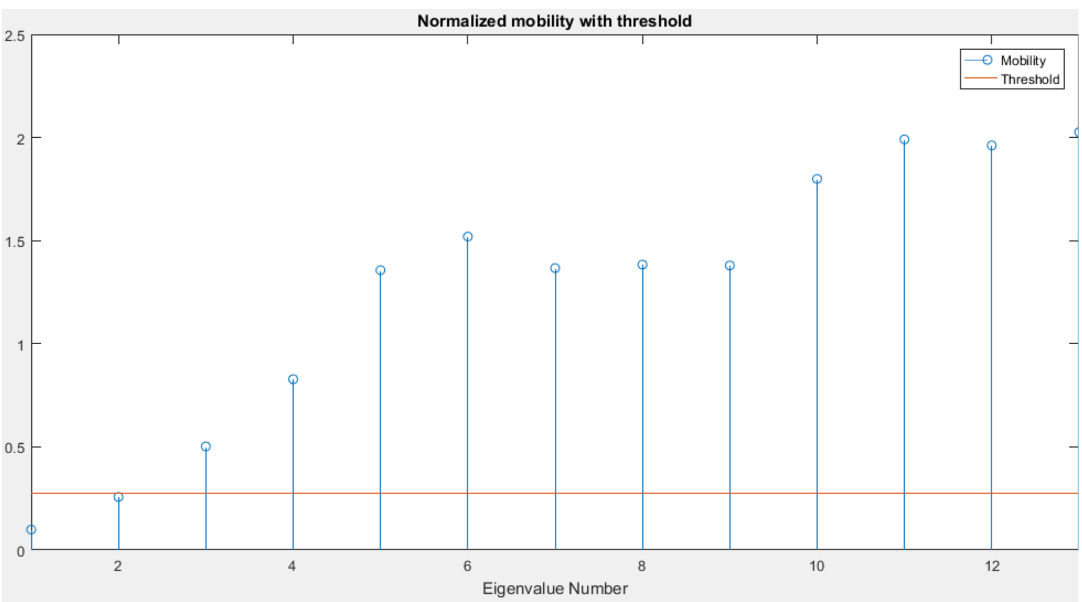

Figure 5. Mobility of eigenvectors of noisy ECG signal

The matrix is then projected onto the eigenvector-created subspace, which is then used to reconstruct the required ECG signal. The eigenvectors that cross the subspace are those whose local fluctuations are smaller compared to the given threshold which is chosen to be 0.275 as per [9]. Finally after the diagonal averaging is done, the required artifact free ECG signal is obtained and shown in Fig.6. Along with this a comparison between noisy ECG signal and extracted ECG signal is shown in Fig. 7

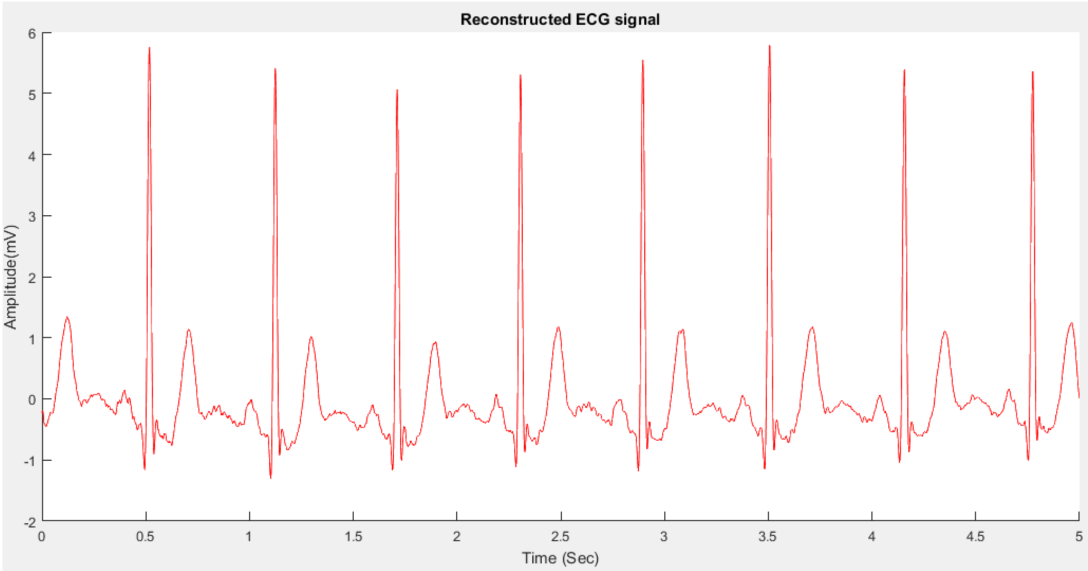

Figure 6. Extracted noise free ECG signal .

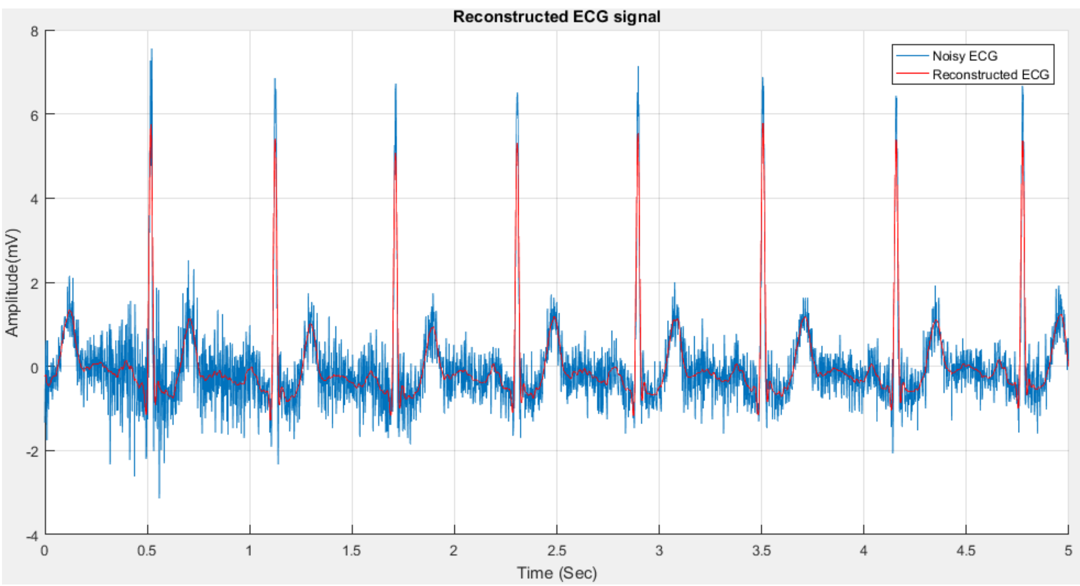

Figure 7. Comparison between noisy ECG signal and extracted ECG signal.

3.3 Comparison with Canonical Correlation Analysis (CCA)

Canonical Correlation Analysis (CCA) is another method to remove the artifacts from ECG signals. Earlier in the research [10] it has been shown that CCA is a better method than SSA when it comes to removal of noise from EEG signals. In this work it has been found that by applying the mobility based grouping technique in SSA along with selecting appropriate window length SSA can show better results than CCA. A comparison between the SSA and CCA method is shown in the table below according to various parameters.

Parameter	Singular Spectrum Analysis	Canonical Correlation Analysis
Correlation Coefficient	0.9826	0.9712
Mean Square Error	0.0385	0.0575
Root Mean Square Error	0.1962	0.2397
Mean Average Error	0.0782	0.1341

4. Conclusion

In order to eliminate the EMG artifacts from the single channel ECG data, we introduced the singular spectrum analysis in this study along with a different grouping technique. The main concept behind this method is that local fluctuations or mobility of the eigenvectors, rather than their magnitudes, are used to group the data in the SSA. The main concept behind this method is that local fluctuations or mobility of the eigenvectors, rather than their magnitudes, are used to group the data in the SSA. With the correlation coefficient being 0.9826 and other error parameters in the table it clearly shows that the suggested strategy effectively gets rid of EMG artifacts while keeping the ECG data intact more efficiently than CCA.

References

[1] S. Chaudhuri, T.D. Pawar and S.P. Duttagupta, "Ambulation analysis in wearable ECG", New York: Springer Science + Business Media, 2009, ch. 1.
[2] G.M. Friesen, T.C. Jannett, M.A. Jadallah, S.L. Yates, S.R. Quint and H.T. Nagle, "A comparison of the noise sensitivity of nine QRS detection algorithms", IEEE Transactions on Biomedical Engineering, vol. 37, no. 1, pp. 85-98, 1990.
[3] P. Kligfield et al., "Recommendations for the standardization and interpretation of the electrocardiogram", Journal of the American College of Cardiology, vol. 49, no. 10, 2007.
[4] F. Shaffer and J.P. Ginsberg, "An overview of heart rate variability metrics and norms", Frontiers in Public Health, vol. 5, no. 258, pp. 1-17, 2017.
[5] https://www.google.com/search?q=ecg+signal&client=opera&hs=uwW&sxsrf=ALiCzsaQdU6Vxkb6Pvy_fo3v7oX3fc3ZQQ:1658715348138&source=lnms&tbm=isch&sa=X&ved=2ahUKEwjXwO-Y_JL5AhWrR2wGHWAcClEQ_AUoAXoECAIQAw#imgrc=PemGHIrJ_HLMTM
[6] A. R. Teixeira, A. M. Tome, E. W. Lang, P. Gruber, and A. M. da Silva, "Automatic removal of high-amplitude artifacts from single channel electroencephalograms," Computer methods and programs in biomedicine, vol. 83, no. 2, pp. 125-138,2006.
[7] K. Najarian and R. Splinter, Biomedical Signal and Image Processing,ser. Biomedical engineering. Taylor & Francis, 2005.
[8] https://archive.physionet.org/cgi-bin/atm/ATM
[9] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 33, no. 2, pp. 387-392, 1985.
[10] Gao J, Zheng C, Wang P. Online removal of muscle artifacts from electroencephalogram signals based on canonical correlation analysis. Clin EEG Neurosci. 2010 Jan;41(1):53-9. doi: 10.1177/155005941004100111. PMID: 20307017.
[11] Klass DW. The continuing challenges of artifacts in the EEG. Am J EEG Technol 1995;35:239–69./
[12] Chen X, Liu A, Peng H, Ward R. A preliminary study of muscular artifact cancellation in single-channel EEG. Sensors 2014;14:18370–89
[13] LiuQ, LiuA, ZhangX, ChenX, QianR, ChenX. Removal of EMG artifacts from multichannel EEG signals using combined singular spectrum analysis and canonical correlation analysis. J Healthc Eng 2019;2019.