10.2478/msr-2013-0001



# Artifact Removal from Biosignal using Fixed Point ICA Algorithm for Pre-processing in Biometric Recognition

Puneet Mishra, Sunil Kumar Singla

Department of Electrical and Instrumentation Engineering, Thapar University, 147004, Patiala, India
puneetmishra04@gmail.com, ssingla@thapar.edu

**In the modern world of automation, biological signals, especially Electroencephalogram (EEG) and Electrocardiogram (ECG), are gaining wide attention as a source of biometric information. Earlier studies have shown that EEG and ECG show versatility with individuals and every individual has distinct EEG and ECG spectrum. EEG (which can be recorded from the scalp due to the effect of millions of neurons) may contain noise signals such as eye blink, eye movement, muscular movement, line noise, etc. Similarly, ECG may contain artifact like line noise, tremor artifacts, baseline wandering, etc. These noise signals are required to be separated from the EEG and ECG signals to obtain the accurate results. This paper proposes a technique for the removal of eye blink artifact from EEG and ECG signal using fixed point or FastICA algorithm of Independent Component Analysis (ICA). For validation, FastICA algorithm has been applied to synthetic signal prepared by adding random noise to the Electrocardiogram (ECG) signal. FastICA algorithm separates the signal into two independent components, i.e. ECG pure and artifact signal. Similarly, the same algorithm has been applied to remove the artifacts (Electrooculogram or eye blink) from the EEG signal**.

**Keywords: Independent component analysis, electroencephalogram, electrocardiogram, electrooculogram**

## 1. INTRODUCTION

ELECTROENCEPHALOGRAM (EEG) signal is a record of the electrical activity of the brain taken from the scalp due to the effect of millions of neurons ($10^9$-$10^{10}$ neurons and $10^{15}$ synapses). The pattern of EEG obtained is very useful for the physician/doctor for the purpose of various diagnostics such as finding of epileptic seizure, tumor and other medical disorders [1]. It has been found with the study that every individual has a different EEG pattern, i.e. EEG is person dependent and differs from individual to individual [2]-[5] and can be used as a source of biometric information. Similarly, studies have been conducted on ECG (which is a measure or interpretation of the electrical activity of the heart) and it has been observed that ECG is different in different individuals, even the twins have different ECG patterns [6], [7] and it can also be used as a source of biometric information.

The special characteristics of EEG and ECG make them a unique tool of human authentication because these characteristics cannot be stolen or copied as those of conventional systems. Only a handful of studies have been done where EEG and ECG signals have been used as a source of biometric information. Few researchers who have worked in the area of EEG have reported a success rate of 70-90% [8]-[14]. Similarly, 90-96% success rate has been reported when ECG has been used as a source of biometric information [6] and [15]-[17].

EEG signals are very much susceptible to artifacts and contain artifacts in the form of eye blink, eye movement, muscular movement, line noise, etc. Blinking of eye produces large electrical potential near the eye known as Electrooculogram (EOG). It is a non-cortical activity which contaminates the EEG while spreading across the scalp. Similarly, ECG signals may contain artifacts in the form of line noise, baseline wandering, tremor artifact, or any random unknown noise due to malfunction in electrodes. These artifacts from ECG and EEG signals must be eliminated to increase the success rate of the system. The traditional ways of artifact removal which include visual inspection of artifacts, notch filters, etc., cannot effectively settle down the problem of filtering the overlapping signals [8]-[12].

A new digital signal processing technique called Independent Component Analysis (ICA), based on multivariate signal statistics, emerged as a novel technique for decomposing the signals into several independent components [19]. Fixed point iteration technique on signals will decompose into several independent components depending on statistical independencies of signals [20]. The method of ICA can be applied to contaminate EEG or ECG signals to separate the noises from the useful signals [20]-[21].

Various techniques have been introduced to improve the performance of the ICA technique. Just like FastICA, Genetic Algorithm based ICA also called GALME-ICA has shown significant results while reducing the EOG artifact from EEG signals in Brain Computer Interface [22]. ICA uses the statistical information to find the hidden independent components from the mixed signals, which may lead to loss of some information, therefore advances in ICA based on Genetic Algorithm have validated on ECG extraction from additive noise and have shown good results [23].

## 2. SUBJECT & METHODS

### 2.1. *Data Collection*

Twenty healthy male volunteers with age ranging between 23-27 years participated in the study. EEG, ECG and EOG signals were recorded for every subject on a three channel data acquisition system (BIOPAC-MP 36, Biopac Systems, Inc.) at a sampling rate of 256 Hz. Subjects were instructed to lie down on the bed with open eyes and relaxed state. EEG was recorded from Fp1 and Fp2 as an active electrode and $C_Z$ as a reference electrode. Similarly, ECG was recorded from the subject as per Einthoven's triangle configuration keeping right leg as reference.

### 2.2. Independent Component Analysis

Independent Component Analysis (ICA) is a method for finding underlying components within a mixture of signals which are both statistically independent and non-Gaussian from multivariate (multi-dimensional) statistical data [24]-[26].

For the ICA model, *n* linear mixtures $x_1, x_2, \dots x_n$ of n independent components combined together in a mixture are denoted as;

$$x_j = a_{j_1} S_1 + a_{j_2} S_2 + \dots\dots + a_{j_n} S_n, \text{ for all } j \quad (1)$$

Whereas, referring to (1) $x$ is a random vector whose elements are a mixture of $x_1, x_2, \dots x_n$ and S be the random vector with components $s_1, s_2, \dots s_n$. The above equation model can be rewritten as the generalized form:

$$X = AS \quad (2)$$

The above model referring to (2) is called Independent Component Analysis or the ICA model. The ICA model is solved with the assumption that components of S are statistically independent, independent components follow non-Gaussian distribution, and the mixing matrix A is square. After estimating the matrix A, we can easily find its inverse by inverse transformation (I), i.e.

$$A^{-1} = I \quad (3)$$

$$S = IX \quad (4)$$

Equation (4) calculates each independent component of S from the mixture of signals. The ICA model very closely resembles the Blind Source Separation (BSS) [27].

### 2.3. FastICA or Fixed Point Iteration Algorithm

There are various measures of finding the non-Gaussianity. In this paper we restrict ourselves to a novel and efficient method of Negentropy for the maximization of Non-Gaussianity by the principle of the FastICA algorithm.

#### 2.3.1. Pre-processing by Centering and Whitening

To make the model zero mean centering is performed. The mathematical expression for centering is shown in (5):

$$x = x - E(x) \quad (5)$$

Similarly, to make the system variance equal to unit, whitening is performed. Whitening reduces the number of parameters to be calculated and hence the complexity. The mathematical expression for whitening is:

$$E\{\hat{x}\hat{x}^T\} = I \quad (6)$$

#### 2.3.2. FastICA for n units

To estimate several independent components, the weights associated with components of **x** have been obtained, i.e. $W_1, W_2 \dots W_N$, as independent sources. Following algorithm has been used for iteration [19], [28]-[29].

(a) Take an initial row vector $Wi$

(b) Apply Newton phase:

$$Wi = E\{\hat{x} g(Wi^T \ \hat{x})\} - E\{g'(Wi^T \hat{x})\} Wi \quad (7)$$

Whereas,

$$g_1(y) = \tanh(a_1 y); \ g_2(y) = y^* e^{(-1/2y^2)}; \quad g_3(y) = 4y^3 \quad (8)$$

(c) Normalization:
Normalize total matrix obtained after the above iterations

$$Wi = (Wi - mean) / std \cdot deviation \quad (9)$$

(d) Decorrelation:

$$Wi = Wi - \sum Wi^T Wj \ Wj \quad (10)$$

(e) Normalization again {Repeat step (c)}

(f) If $W^T(i) * W(i-1)$ is not close enough to 1,

Let $W_{i+1}$, and go back to step (b) (11)

## 3. RESULTS

All the computations have been made in MATLAB R2009a (Propriety of Matrix laboratory). As explained in earlier sections, eye-blink and EEG signals are generated by different sources that are independent from each other. Thus, we can use FastICA to separate the signals, i.e. eye-blink artifact and EEG into statistically independent components. A similar operation has been carried out where random noise signal and ECG are generated at two different sources that are different from each other. After the iterations using fixed point algorithm application of all, the desired artifact free signal has been obtained for further evaluation.

### 3.1. Application of ICA on ECG signal

ICA has been used to extract independent components from ECG signal which is synthetically mixed with the random noise as shown in Fig.1(a). Recording has been made in Biopac MP-36 Data acquisition card. It can be seen from Fig.1(b) that ECG signal from 0 to 2000 ms is highly contaminated with random noise, which must be removed to obtain the pure ECG signal. The concept of two independent components, i.e. one is ECG pure signal and the other as noise, has been applied. The FastICA algorithm when applied on the signal of Fig.1(b) efficiently decomposed the signal into its individual components as shown in Fig.1(c) and Fig.1(d).

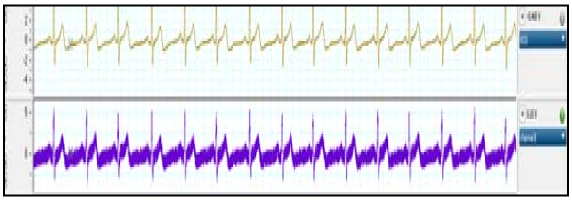

Fig.1(a). ECG signal mixed with random noise signal in Biopac MP-36

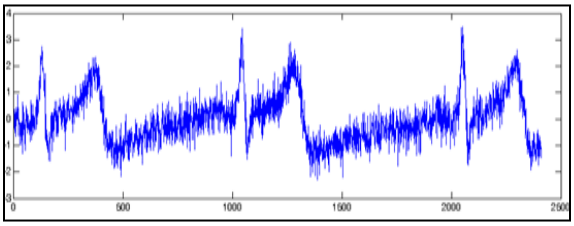

Fig.1(b). Contaminated ECG component for processing in Matlab

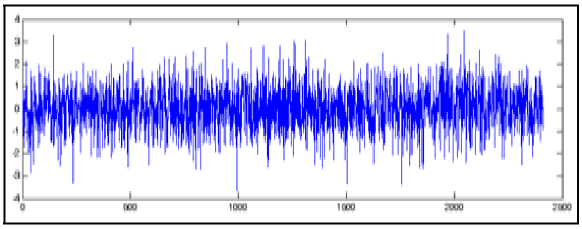

Fig.1(c). Extracted noise from ECG signal using FastICA

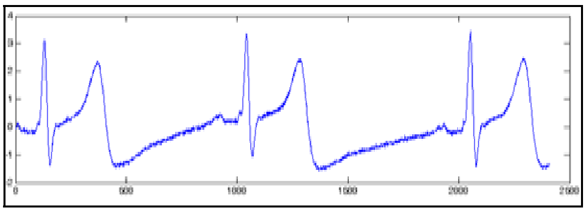

Fig.1(d). Pure ECG signal retrieved using FastICA

### *3.2. Application of ICA on EEG signal*

The recorded EEG signal from subjects, which was contaminated with eye blink artifact, is shown in Fig.2(a). As can be seen from Fig.2(b), the portion of signal from 1000-1600 ms is showing high amplitude variation due to eye blink. This is the case of origin of application of ICA where two independent components, i.e. EEG and eye blink are mixed together. The application of FastICA algorithm to the mixed signals lead to the separation of EEG and eye blink artifacts as shown in Fig.2(c) and Fig.2(d), respectively.

Signal to Noise Ratio of the output signal, i.e. after application of ICA iterations has been tabulated under Table 1. SNR is a technical term used to characterize the quality of the signal to the background or noise signal [30]. The SNR value for the EOG extracted from EEG (ID_001 to ID_020) is shown in Table 1, including all other 19 subjects. The mathematical expression to calculate SNR in *dB* is:

$$SNR = 20\log_{10}(E_S / E_N) \qquad (12)$$

From Table 1, it has been concluded that SNR has shown a significant change from lower to slightly higher values, which shows a change in signal purification due to ICA implementation. Some loss of information may occur while undergoing visual inspection, because it is an identity of Blind Source Separation.

The application of ICA is not limited to signal processing but also extends from separation of mixed voices and images, analysis of various types of data and data extraction in financial matters, data communication, system identification and several other biomedical signal processing tasks [24], [31].

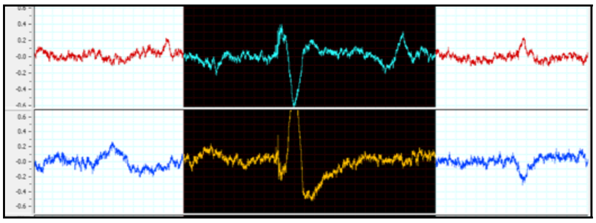

Fig.2(a). EEG signal with eye blink as recorded in Biopac MP-36

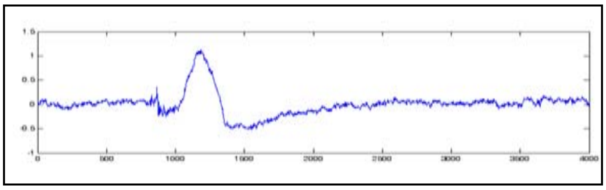

Fig.2(b). Portion of EEG contaminated with eye blink

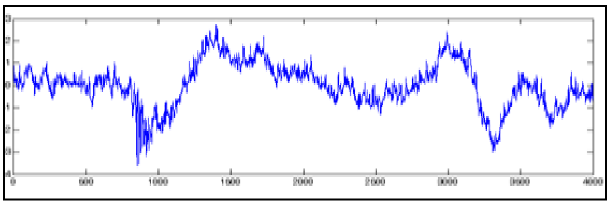

Fig.2(c). Artifact free EEG signal after FastICA application

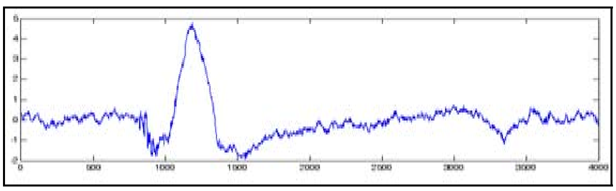

Fig.2(d). Eye blink artifact separated out after FastICA

Table 1. Signal to Noise Ratio (SNR) of EEG signals after ICA iterations.

| SUBJECT_ ID | Signal to Noise Ratio after ICA iterations |
|---|---|
| ID_001 | 6.5062 |
| ID_002 | 4.4344 |
| ID_003 | 4.1374 |
| ID_004 | 5.1663 |
| ID_005 | 5.4280 |
| ID_006 | 4.6987 |
| ID_007 | 5.0948 |
| ID_008 | 4.6073 |
| ID_009 | 5.8285 |
| ID_010 | 5.9334 |
| ID_011 | 6.1028 |
| ID_012 | 5.1193 |
| ID_013 | 4.6500 |
| ID_014 | 5.2688 |
| ID_015 | 4.8219 |
| ID_016 | 4.5228 |
| ID_017 | 5.6289 |
| ID_018 | 5.0124 |
| ID_019 | 5.4578 |
| ID_020 | 4.2933 |

## 4. DISCUSSION / CONCLUSIONS

The results presented in this paper appear to be promising. The ICA is a powerful tool for extracting the independent sources from the EEG and ECG signal (mixed with artifacts) with the concept that components are statistically independent. The efficacy of this algorithm has been proven with synthetic as well as real EEG signal. ICA is a fairly novel technique and found useful while being applied to the biomedical signal processing tasks. The approach mentioned here will rectify the traditional method of artifact removal while evaluating EEG or ECG for biometric authentication. This technique will lead to better results while interfacing it as a pre-processing step during biometric recognition. ICA could also be used for extraction of other features in financial data or biomedical signals.

## REFERENCES


[1] Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review*, 2 (2), 1-11.

[2] Eischen, S.E., Luckeritz, J.Y., Polish, J. (1995). Spectral analysis of EEG from families. *Biological Psychology*, 41, 61-68.

[3] Svidevskaya, N.E., Koroľkova, T.A. (1995). Genetic features of the human cerebral cortex. *Neuroscience and Behavioural Physiology*, 25 (5), 370-376.

[4] Vogel, F. (1970). The genetic basis of the normal EEG. *Human Genetics*, 10, 91-114.

[5] Poulos, M., Rangoussi, M., Hrissikopoulos, V., Evangelou, A. (1999). Person identification based on parametric processing of the EEG. In *Proceedings of the IEEE International Conference on Electronics, Circuits, and Systems*, 5-8 September 1999. IEEE, 283-286.

[6] Biel, L., Pettersson, O., Philipson, L., Wide, P. (2001). ECG analysis: A new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*, 50, 808-812.

[7] Irvine, J.M., Wiederhold, B.K., Gavshon, L.W., Israel, S., McGehee, S.B., Meyer, R., Wiederhold, M.D. (2001). Heart rate variability: A new biometric for human identification. In *International Conference on Artificial Intelligence (IC-AI 2001).* Las Vegas, Nevada, 1106-1111.

[8] Gupta, C.N., Palaniappan, R., Swaminathan, S. (2008). On the analysis of various techniques for a novel Biometric system. *International Journal of Medical Engineering and* Informatics, 1 (2), 266-273.

[9] Ursulean, R., Lazar, A.M. (2009). Detrended cross-correlation analysis of biometric signals used in a new authentication method. *Electronics and Electrical Engineering*, 1 (89), 55-58.

[10] Gupta, C.N., Khan, Y.U., Palaniappan, R., Sepulveda, F. (2009). Wavelet framework for improved target detection in oddball paradigms using P300 and gamma band analysis. *Biomedical Soft Computing and Human Sciences*, 14 (2), 61-67.

[11] Palaniappan, R., Eswaran, C. (2009). Using genetic algorithm to select the presentation order of training patterns that improves simplified fuzzy ARTMAP classification performance. *Applied Soft Computing*, 9, 100-106.

[12] Abdullah, M.K., Subari, K.S., Cheang Loong, J.L., Ahmad, N.N. (2010). Analysis of the EEG signal for a practical biometric system. *World Academy of Science, Engineering and Technology*, 68, 1123-1127.

[13] Yang, X., Dai, J., Zhang, H., Wu, B., Su, Y., Chen, W., Zheng, X. (2011). P300 Wave based person identification using LVQ neural network. *Journal of Convergence Information Technology*, 6 (3), 296-302.

[14] Gupta, C.N., Palaniappan, R., Paramesran, R. (2012). Exploiting the P300 paradigm for cognitive biometrics. *International Journal of Cognitive Biometrics*, 1 (1), 26-38.

[15] Singh, Y.N., Singh, S.K. (2012). Evaluation of electrocardiogram for biometric authentication. *Journal of Information Security*, 12 (3), 39-48.

[16] Israel, S.A., Scruggs, W.T., Worck, W.J., Irvine, J.M. (2003). Fusing face and ECG for person identification. In *Proceedings of IEEE Applied Imagery Pattern*

*Recognition Workshop*, 15-17 October 2003. IEEE, 225-231.

[17] Shen, T.W., Tompkins, W.J., Hu, Y.H (2002). One-lead ECG for identity verification. In *Proceedings IEEE EMBS/BMES Conference*, 23-26 October 2002. IEEE, 62-63.

[18] Makeig, S., Bell, A.J., Jung, T.P., Sejnowski, T.J. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems*, 8, 145-151.

[19] Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10 (3), 626-634.

[20] Vigario, R., Sarela, J., Jousmaki, V., Hamalainen, M., Oja, E. (2000). Independent component approach to the analysis of EEG and MEG recordings. *IEEE Transactions on Biomedical Engineering*, 47 (5), 589-593.

[21] Jiang, J.A., Chao, C.F., Chiu, M.J., Lee, R.G., Tseng, C.L., Lin, R. (2007). An automatic analysis method for detecting and eliminating ECG artifacts in EEG. *Computers in Biology and Medicine*, 37, 1660-1671.

[22] Gupta, C.N., Palaniappan, R. (2011). Reducing power spectral density of eye blink artifact through improved genetic algorithm. In *International Conference on Bioinformatics and Biomedical Technology*, 25-27 March 2011.

[23] Palaniappan, R., Gupta, C.N. (2006). Genetic algorithm based independent component analysis to separate noise from electrocardiogram signals. In *Proceedings of IEEE International Conference on Engineering of Intelligent Systems*, 22-23 April 2006. IEEE, 1-5.

[24] Hyvarinen, A., Erkki, O. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, 13 (4-5), 411-430.

[25] Ungureanu, M., Bigan, C., Strungaru, R., Lazarescu, V. (2004). Independent component analysis applied in biomedical signal processing. *Measurement Science Review*, 4, 1-8.

[26] Krishnaveni, V., Jayaraman, S., Kumar, P.M., Shivakumar, K., Ramadoss, K. (2005). Comparison of independent component analysis algorithms for removal of ocular artifacts from electroencephalogram. *Measurement Science Review*, 5 (2), 67-78.

[27] Jung, T.P., Makeigh, S., Humphries, C., Lee, T.W., Mckeown, M.J., Iragui, V., Sejnowski, J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37, 163-178.

[28] Xu, L., Cheung, C., Yang, H., Amari, S. (1997). Independent component analysis by the information-theoretic approach with mixture of densities. *IEEE Transactions on Neural Networks*, 5, 1821-1826.

[29] Agrawal, G., Singh, M., Singh, V.R., Singh, H.R. (2008). Reduction of artefacts in 12-channel ECG signals using FastICA algorithm. *Journal of Scientific and Industrial Research*, 67, 43-48.

[30] Palaniappan, R. (2010). *Biological Signal Analysis.* Ventus Publishing.

[31] Patil, D., Das, N., Routray, A. (2011). Implementation of fast-ICA: A performance based comparison between floating points, and fixed point DSP platform. *Measurement Science Review*, 11 (4), 118-124.

Received April 4. 2012.
Accepted January 28, 2013.