Feature Extraction Using Independent Component Analysis Method from Non-Invasive Recordings of Electroencephalography (EEG) Brain Signals

Muhammad Azhar[1,2], Ishfaque Ahmed[1,4], Syed Tanveer Iqbal[1,7], Muhammad Jahangir[3,5], Rizwan-ur-Rehman[3,5], Nasir Ali Shah[3,6] and Imran Siddiqui[3,*]

[1]*NED University of Engineering & Technology, Karachi, Pakistan*

[2]*Atomic Energy Medical Centre, JPMC Karachi, Pakistan*

[3]*Department of Physics, University of Karachi, Karachi, Pakistan*

[4]*Government Degree Science College, Gambat, Pakistan*

[5]*Government National College No.1, Shaheed-e-Millat Road Karachi, Pakistan*

[6]*Government Degree College Stadium Road S.R.E. Majeed Karachi, Pakistan*

[7]*DJ Science College Karachi, Karachi, Pakistan*

Abstract: Electroencephalography (EEG) is a well known procedure in neuroscience, performed to extract brain signal activity associated with voluntary and involuntary tasks. Scientists and researchers working in neuroscience are involved in the research of brain computer interfacing (BCI) and in improving the existing BCI systems. In BCI, it is possible for a person to control the external devices remotely using brain signals without neurophysical intervention. In the proposed work the new algorithm is introduced to extract the feature from EEG based recorded brain signals. The features are extracted for a specific motoryaction that is raising the right hand. The proposed algorithm is also verified from EEGLAB routines also based on Independent Component Analysis (ICA) method written in MATLAB platform.

Keywords: BCI, Electroencephalography, ERP, ICA, Infomax, SOBI, FastICA.

1. INTRODUCTION

Brain-Computer Interface (BCI) is the pathway to communicate between the human brain and an external or artificial device. The purpose of the BCI is to detect and quantify characteristics of brain signals. The brain signals are basically the electrical activity of brain, generated due to imbalanced concentration of ions across the membrane of neuron (lipid bi-layer) [1]. To achieve BCI, different methods are used. Most of them are based on EEG recorded from scalp. The EEG data is recorded while the user imagines different things (for example, moving right hand). Motor imagery is a mental strategy used to operate BCI [2]. The mental cursor dominantly controlled by mu rhythms and beta rhythms [3].

Electroencephalography (EEG) is one of them which is most widely used to examine the brain activity. EEG is differ from the other tools that provide the dynamic information of the brain or even indicate whether the patient is asleep or awake [4]. When two (or more) electrodes are positioned on the scalp, the biopotential can be measured between them. The results of electroencephalogram obtained by summation of postsynaptic potentials. The biopotential is about 10-100μV in amplitude because only a small amount (fraction) of pyramidal cell current penetrates the meningeal coverings, cerebral spinal fluid (CSF), and skull to the scalp [5]. A fundamental signal feature is the measure of potential (appeared between a pair of electrodes) at a particular time for a particular event or stimulus. The event related potentials (ERPs) are very small voltages that are generated in the brain in response to specific event such as hand raise, grasping some object, eye blinking etc. ERPs can be extracted by a variety of cognitive, sensory or motor events. In humans, the ERPs can be divides into two categories, sensory and cognitive. ERP waveforms are termed as P50, N100 or N1 wave, P200 or P2 wave, N200 or N2 wave, N2a, N2b, N2c, N300, P300, N400, P600 [6]. P300 is appeared as the most important and the most studied component of the ERP [7]. During recording EEG, some unwanted signals are also generated which cannot be controlled because they are caused by uncontrolled activities like eye blinking, breathing, or heart beating etc. these unwanted signals are called artifacts [8]. There are several techniques

*Address correspondence to this author at the Department of Physics, University of Karachi, Karachi, Pakistan; Tel: 0092(0)335 347 2887;
E-mail: imranahmadsiddiqui@yahoo.com

PACS numbers: 8435, 8780

ISSN: 1814-8085 / E-ISSN: 1927-5129/17

adopted by researchers to decontaminate the signals from artifacts. In this research work ICA implemented to extract the features and to remove the artifacts.

Independent component analysis (ICA) is a useful technique helps to find components from multidimensional statistical data. The ICA looks for components that are statistically independent, and nongaussian. This characteristic makes it different from other methods [9].

The ICA method is based on the simple physically realistic assumption (in case of brain signals even in any other phenomenon) that if different signals are generated through different physical processes then those signal must be mathematically independent. If this assumption is reversed then it leads to a new assumption (logically unwarranted but which works in practice), namely: if signal which are statistically independent are extracted from a mixture of signals then the signals must be related from different physical processes. Accordingly, ICA separates the mixed signals into unmixed signals which are statistically independent. If the above assumption is valid then then each of the signals extracted by ICA will have been produce (or generated) by an unrelated process and the will be a desired signal [10].

Radüntz *et al*. (2015) proposed an automated classifier based on ICA, compare its results with manual expert inspection and report 88% accuracy rate of classifier. Classifier has advantages that it is independent of type and number of artifacts [11].

A temporal ICA model was introduced by Eichele *et al*. (2011) for event related EEGs. The accuracy of the analysis for decomposition as well as re-construction depends on inter-individual phase locking among the event related EEG signals [12].

Bou *et al*. (2014) reported that ICA is one of the most effective artifact rejecting technique for various types of artifacts. Automated ocular artifact rejection based on ICA was studied without additional channels. Ocular ICs were then classified by using Kmeans [13].

De *et al*. (2011) introduced a new algorithm which is capable to incorporate with Jade (Joint Approximation Diagonalization Eigenmatrices) and Sobi (second Order Blind Identification) implementations of ICA. Algorithm is based on assumption of known mixing matrix whether it is single or more [14].

Naeem M *et al*. (2006) recorded the EEG data by placing 22 electrodes over the scalp during motor imagery tasks that consist of four different classes, namely the imagination of tongue, right hand, left hand and foot movements. They have recorded two different sessions for eight subjects and used three different ICA algorithms (Infomax, SOBI and FastICA) to compare common spatial patterns (CSP). They found algorithm (Infomax) as best performance algorithm [15].

Milanesi *et al*. (2008) modified a form of FastICA algorithm and applied it to deal with convolutive mixtures. The key idea of this algorithm is that convolution changes into linear mixing in frequency domain and ICA could estimate sources as usually. They used FastICA for dealing with complex values. The paper shows interesting results that are obtained by this algorithm to 9 EEG recordings [16].

2. METHODOLOGY

2.1. Subject and Experimental Paradigm

A healthy female person participated in the experiment, here is termed as the subject. Prior to the actual EEG data acquisition, a written consent was also taken from the subject. A proper protocol regarding the data acquisition was first briefed to the subject and was initially trained to properly acquire brain data. The experiment was performed for recording EEG signals associated with the activity corresponding to the raising of right hand, herein termed as activity PULL. During the experiment the subject was advised to sit on an armchair in front of which a monitor screen was fixated. On the screen a queue suggesting the subject to raise her right hand that is PULL message is displayed.

2.2. EEG Recording

EEG data was collected by placing 14 electrodes (AF3 · F7 · F3 · FC5 · T7 · P7 · O1 · O2 · P8 · T8 · FC6 · F4 · F8 · AF4) on the subject's scalp. The alphabets in the name of each electrode identify the lobe (F=Frontal lobe, T=Temporal lobe, C=Central lobe, P=Parietal lobe, O=Occipital lobe) and the numeric values identify the hemisphere location. A single EPOC contains 8064 data points correspond to 14 electrodes. Data was collected at sampling rate 128Hz. It is also known as srate.

10-20 system was used to locate the channel position. 10-20 system is an internationally recognized system that provide the uniform coverage of the skull. In this system, the surface area of scalp is divided into 10% or 20%. The basic purpose of this system is to locate the electrodes in such a way that distance

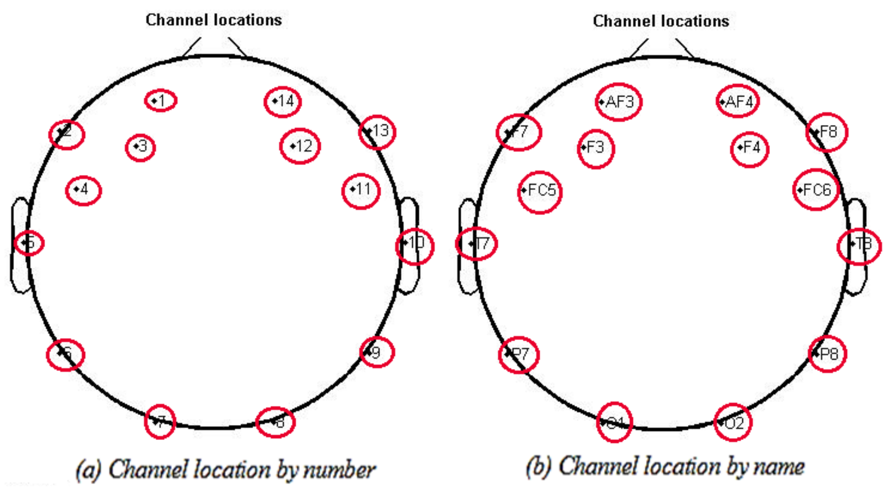

Figure 1: Electrodes location.

between adjacent electrodes remains either 10% or 20% of total surface area. The odd electrodes (AF3 · F7 · F3 · FC5 · T7 · P7 · O1) were placed on left side and even ones (O2 · P8 · T8 · FC6 · F4 · F8 · AF4) were placed on the right side of the skull [17]. The Figure **1** shows channel location using 10-20 system.

2.3. Signal Processing

Brain signal processing allows one to understand, interpret, and decode brain signals. Before analyzing brain signals, they need to be appropriately processed, for example, to remove artifacts. The specific frequency band is also obtained by signal processing. The proper study of functionalities of brain for various purpose are based on the well processed signals.

In this research brain signals are processed using two techniques:

- EEGLAB
- ICA algorithm

2.3.1. EEGLAB

EEGLAB is widely used software in analysis of brain signals and in visualization of brain dynamics. In this research work EEGLAB software (EEGLABv11.0.3.1b) used to process the EEG data. In the first step EEGLAB was run by using MATLAB. The original data and emotive file (contained 14 channel EEG information) was feed to EEGLAB. Channel location was also found by this software. The data was then filtered (in order to remove the linear trend) by using basic FIR filter in which signals are filtered by Lower edge frequency band pass as 5Hz and higher edge frequency band pass as 40 Hz. The filtered data then was run using run ICA function. Finally results were obtained.

2.3.2. Modified ICA algorithm

The data contains the signals collected from 14 channel device. A single epoch of the data consist of 8064 data points corresponds to 14 channels. A matrix "X" of order 8064x14 was obtained.

$$X = \begin{bmatrix} x_{11} & x_{12} & . & . & x_{1n} \\ x_{21} & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{m1} & . & . & . & x_{mn} \end{bmatrix} \tag{1}$$

Where

m = 1, 2, 3,… , 8064

n = 1, 2, 3, … , 14

Here x_n is the voltage signal recorded at *nth* electrode. In this single electrode the signals of other neighboring electrodes also contribute. That is why this matrix is also known as mixing matrix.

In order to reduce the number of rows in matrix "X" moving average applied to obtain matrix "X_{avg}" of order 14x14.

$$X_{avg} = \begin{bmatrix} x_{11} & x_{12} & . & . & x_{1n} \\ x_{21} & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{m1} & . & . & . & x_{mn} \end{bmatrix} \tag{2}$$

Where

m = 1, 2, 3,… , 14

n = 1, 2, 3, … , 14

An unknown matrix “S” was considered as:

$$S = \begin{bmatrix} s_{11} & s_{12} & . & . & s_{1n} \\ s_{21} & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ s_{m1} & . & . & . & s_{mn} \end{bmatrix} \quad (3)$$

Where, S_{mn} show separated signal (component) of *mth* location recorded by *nth* electrode.

The mixing matrix “X” must be written in terms of unmixing matrix “S” if matrix “S” is multiplied by another unknown mixing matrix “A”.

$$X = AS \quad (4)$$

From the above equation matrix “S” can be written as:

$$S = A^{-1}X \quad (5)$$

Let $A^{-1} = W$

$$S = WX \quad (6)$$

In the above equation, there are two unknown matrices, matrix “S” and matrix “W”. Here the ICA algorithm was used to find the matrix “W”. Matrix “W” is nothing but the inverse of matrix “A”. The matrix “A” was calculated from matrix “X” with the help of coefficient of determination “r^2” (square of Karl Pearson’s coefficient of correlation “r”).

$$r_{xy} = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}} \quad (7)$$

Where r_{xy} represent the coefficient of correlation between signals generated on electrode “x” and electrode “y”.

Matrix “A” was introduced as:

$$A = \begin{bmatrix} a_{11} & a_{12} & . & . & a_{1n} \\ a_{21} & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{m1} & . & . & . & a_{mn} \end{bmatrix} \quad (8)$$

Where $a_{mn} = 1 - r_{mn}^2$

The aim of this modified ICA algorithm is to find out unmixing matrix “W” which is the inverse of mixing matrix “A”. Therefore, each component of matrix “A” must be uncorrelated. As the coefficient of determination “r_{mn}^2” gives the maximum degree of correlation so the measure of un-correlation can be written as “$1 - r_{mn}^2$”.

2.3.3. Feature Extraction and Classification

The components and feature sets, extracted by EEGLAB and modified ICA algorithm were examined with the help of power spectrum analysis [18] and topographic maps [19]. Both applied techniques then compared by visual graphical analysis. The visual examination of power spectrum for channel properties led us to conclude that 1 out of the 14 channels represented the task-related activity. The similar conclusion was obtained from power spectrum of component properties.

3. RESULTS AND DISCUSSION

This section discusses the results obtained from the proposed methodology. Data is processed by using MATLAB assisted software, EEGLAB and the algorithm based on the principles of independent component analysis (ICA). The results obtained from both techniques compared to draw the conclusion.

3.1. Results from EEGLAB-Power Spectrum Analysis

Topographic information were obtained with the help of power spectrum analysis. Power spectrum is a quantitative analysis of electroencephalogram (EEG) that based on fast Fourier transform with the help of computer software. Power spectrum analysis can be used to compare different EEG statistically [18]. Spectral plot of power spectrum represents the spectrum of the activity of one data channel. It provide a way to represent the magnitude of signals at measurement points with colors. The power spectrum was the plotted between log of power (measured in $\mu V^2/Hz$) vs frequency (measured in Hz).

As the motor imagery containing mu or alpha rhythm (7Hzto 13Hz) that’s why these frequency signals were specially analyzed through power spectrum to extract the feature and to find where the feature is decomposed.

Figures **2** and **3** represent the power spectrum of the recorded EEG data on EEG electrodes and power

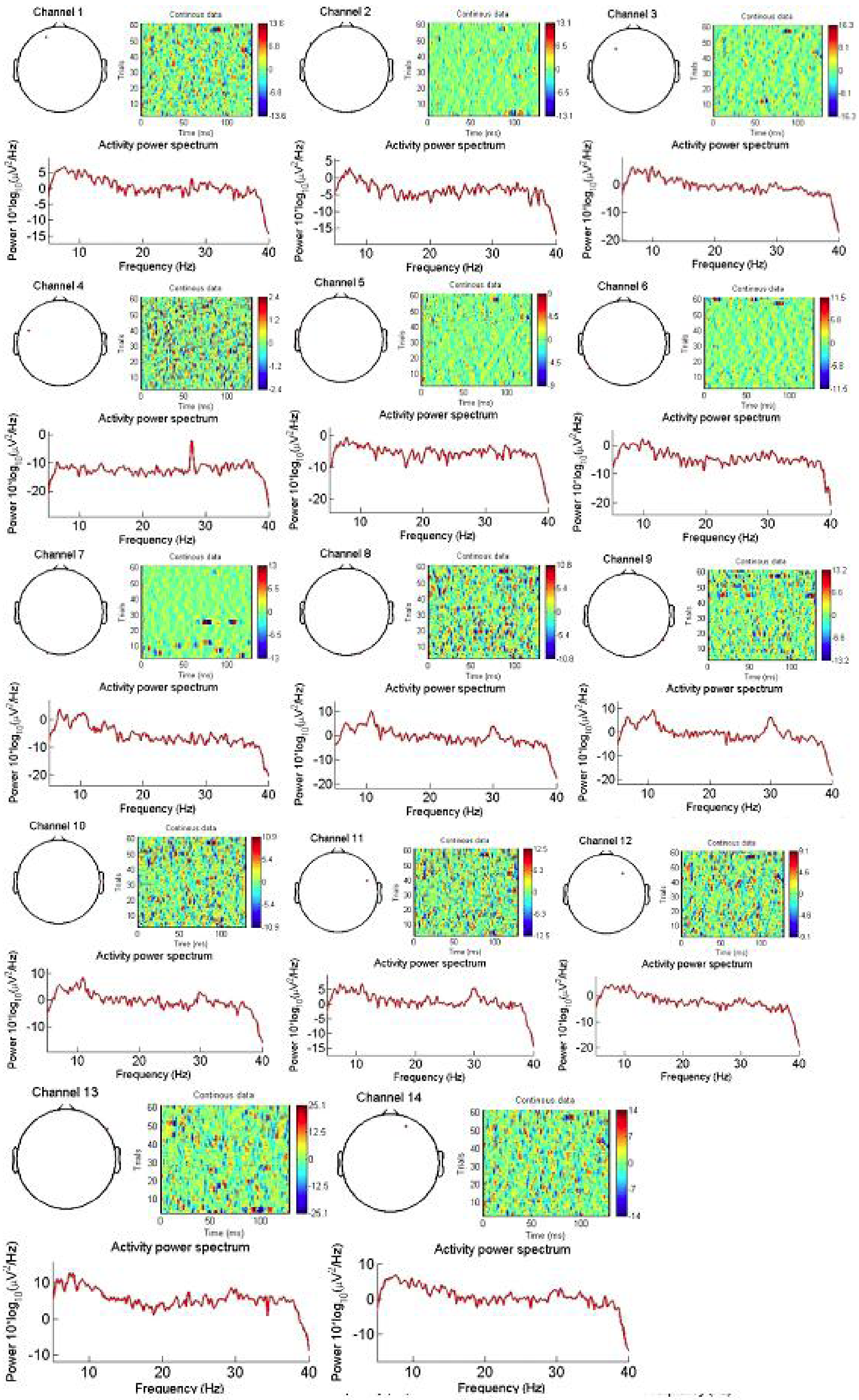

Figure 2: Channel properties through power spectrum.

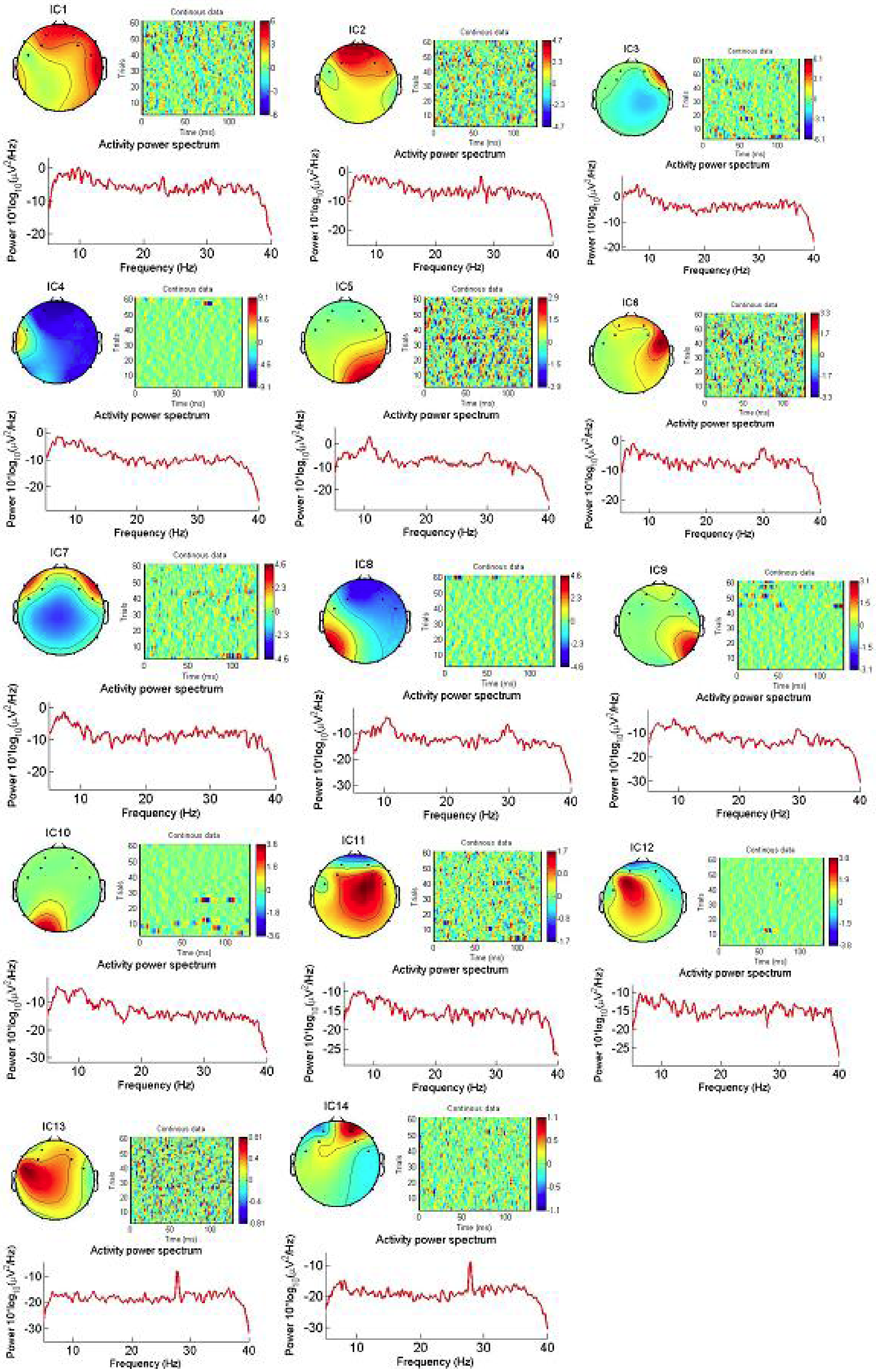

Figure 3: Component properties through power spectrum.

spectrum of the decomposed independent components respectively. Power spectrum of signal recorded on each electrode and each independent component is separately calculated via EEGLAB.

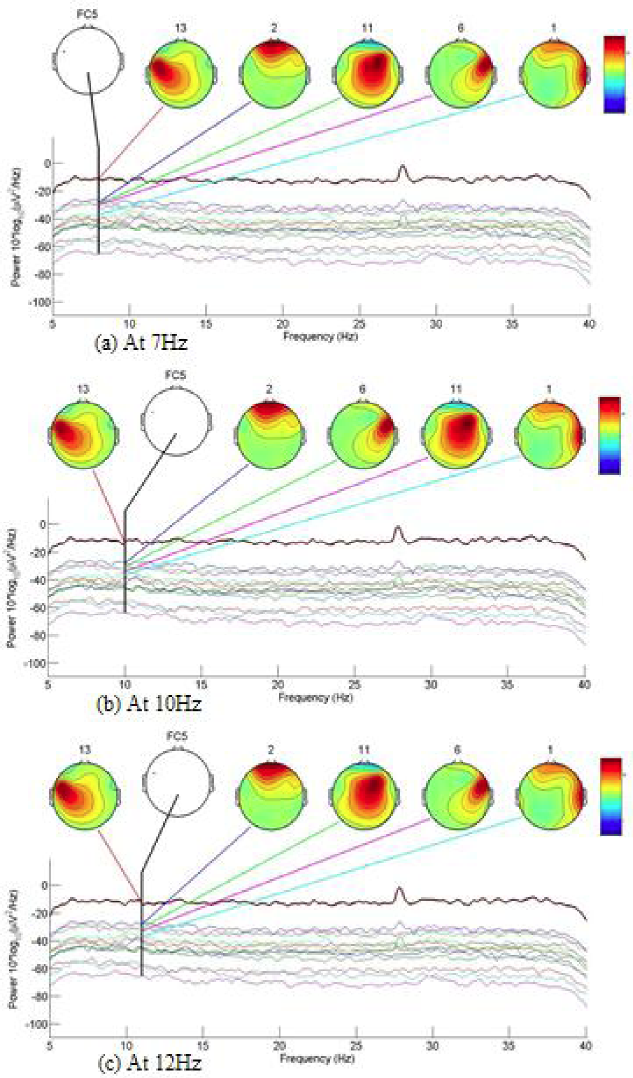

Figure 4: Component spectra and map at 7Hz, 10Hz and 12 Hz.

In Figure **5**, 14 ICs were chosen to depict the motor imagery task for right hand raise. These 14 ICs were then compared to on the basis of topographic information. In these topographic images the powers of signals are distributed on the basis color. In which the intensity of power is ranges from -1(represents by dark blue color) to + 1 (represents dark red).

3.2. Results From Modified ICA Algorithm

The Figures **6**, **7** and **8** show the results obtained from modified ICA algorithm. These figures represent the identification of independent component and the channel at which the activated signal (feature) is decomposed.

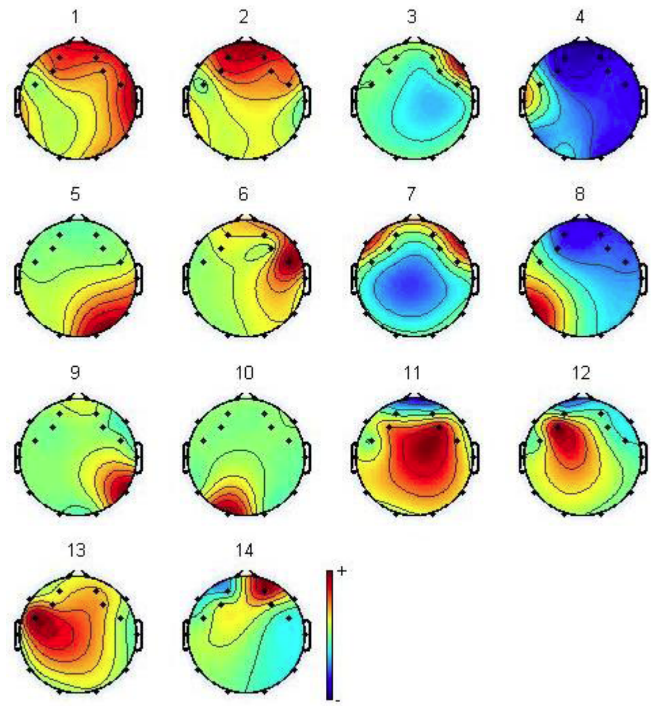

Figure 5: 14 IC spatial distributions in a representative subject (blue represents –1, red represents 1).

4. DISCUSSION AND CONCLUSION

The actual data of order 8064×14 was processed using EEGLAB. The Figures **2** and **3** shows the channel properties and component properties obtained by the graph between log of power $(\mu V^2/Hz)$ vs frequency (Hz) A fine peak in power spectrum is observed in Figure **2** which is obtained by channel 4. The similar peak of similar trend is observed in Figure **3** which is obtained by IC13.

As our desired signal are basically motor signal that contains alpha/mu rhythm (ranges from 7 to 13Hz) so the signals was filtered using FIR filter to remove the linear trend of the signals. Figure **4a**, **4b** and **4c** show the power spectrum at frequencies (frequencies with maximum power) 7Hz, 10Hz and 12Hz respectively. The results from Figure **4a**, **4b** and **4c** show that desired signal is generated at 4^{th} channel left frontal central (FC5) and the signal is decomposed at component 13 (IC-13).

The Figure **5** shows the component map in 2 dimensions. The topographic Image of component 13 shows the best contrast. Dark red color shows the origin of activated signal which is the position of FC5 electrode. The overall analysis shows that the feature is generated at channel 4 (FC5) while the decomposing of signal occurs at component 13.

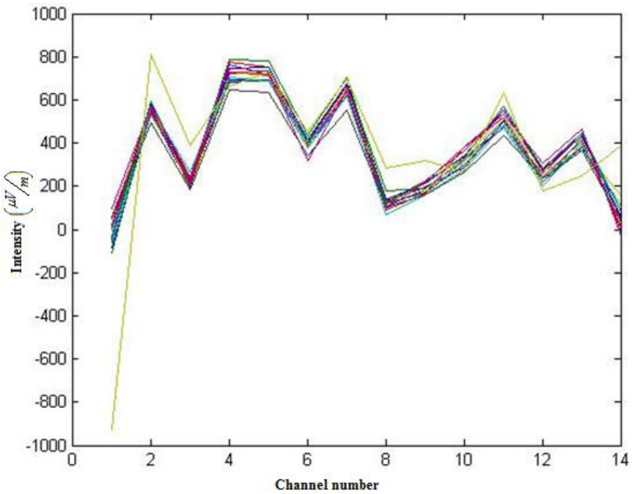

Figure 6: The Line Plot representing the consistency & signal strength of components.

The graphs are formed from matrix "S". Figure **6** shows the line plot, plotted between intensity of the signals and channel number in which channel 4 appeared to be more consistent.

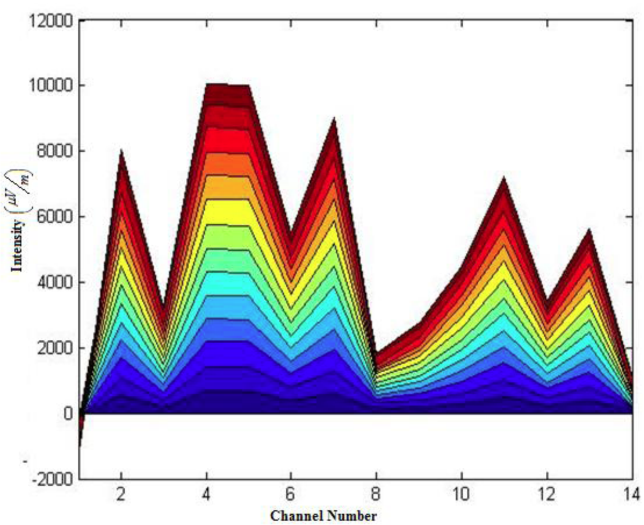

Figure 7: Area plot represents the signal strength.

Figure **7** represents the area obtained by graph Plotted between signal strength (10 times of intensity) vs channel number which shows the similar result as Figure **6**.

The Figure **8** represents the imgesc figure shows the results by plotting image between channel number and sample number (IC) in which channel 4 and IC-13 having the dark color as compare to others that is the evidence existence of activated signal.

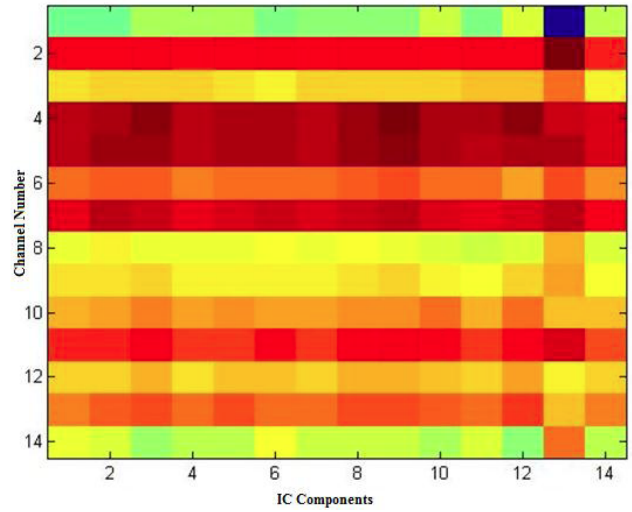

Figure 8: Imagesc represent the activity recorded on channel.

4.1. Comparison between Proposed Techniques

The results drawn from two different techniques (EEGLAB and modified ICA algorithm) adopted in this research found to be correlated. Brain signals are decomposed and features are efficiently extracted by these two methods.

5. CONCLUSION

The brain signals are recorded at the surface of scalp using 14 channel EEG devices. The artifacts and noises are associated with recorded data. The data is processed in order to remove the artifacts and noises. Features are extracted with the help of proposed ICA algorithm which is based on coefficient of determination.

The proposed ICA algorithm found to be an efficient tool for brain signals feature extraction. Results are obtained and then compared. The proposed algorithm is verified by EEGLAB.

REFERENCE

[1] Rajesh P. Rao N. Brain-Computer Interfacing. Cambridge University Press 2013; p. 2.

[2] Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan T. Brain-computer interfaces for communication and control. Clin Neurophysiol 2002; 113: 767-91. https://doi.org/10.1016/S1388-2457(02)00057-3

[3] Cheng M, Jia W, Gao X, Gao S, Yang F. Mu rhythm-based cursor control: an offline analysis. Clin Neurophysiol 2004; 115: 745-51. https://doi.org/10.1016/j.clinph.2003.11.038

[4] Libenson MH. Practical Approach to Electroencephalography, Elsevier Health Sciences 2012.

[5] Baura GD. Medical Device Technologies: A Systems Based Overview Using Engineering Standards, Academic Press 2012.

[6] Sur S, Sinha VK. Event-related potential: An overview. Industrial Psychiatry Journal 2009; 18(1): 70. https://doi.org/10.4103/0972-6748.57865

[7] Fazel-Rezai R, Ramanna S. Brain Signals: Feature Extraction and Classification Using Rough Set Methods, Lecture Notes in Computer Science 2005; pp. 709-718. https://doi.org/10.1007/11548706_75

[8] Shoker L, Sanei S, Chambers J. Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm. Signal Processing Letters IEEE 2005; 12(10): 721-724. https://doi.org/10.1109/LSP.2005.855539

[9] Hyvärinen A, Karhunen J, Oja E. Independent Component Analysis, John Wiley & Sons 2004.

[10] Stone JV. Independent Component Analysis: A Tutorial Introduction, MIT Press 2004.

[11] Radüntz T, Scouten J, Hochmuth O, Meffert B. EEG artifact elimination by extraction of ICA-component features using image processing algorithms. Journal of Neuroscience Methods 2015; 243: 84-93. https://doi.org/10.1016/j.jneumeth.2015.01.030

[12] Eichele T, Rachakonda S, Brakedal B, Eikeland R, Calhoun VD. EEGIFT: group independent component analysis for event-related EEG data. Computational Intelligence and Neuroscience 2011; 9.

[13] Bou Assi E, Rihana S, Sawan M. Kmeans-ICA based automatic method for ocular artifacts removal in a motorimagery classification. In Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE 2014; pp. 6655-6658.

[14] De Vos M, De Lathauwer L, Van Huffel S. Spatially constrained ICA algorithm with an application in EEG processing. Signal Processing 2011; 91(8): 1963-1972. https://doi.org/10.1016/j.sigpro.2011.02.019

[15] Naeem M, Brunner C, Leeb R, Graimann B, Pfurtscheller G. Seperability of four-class motor imagery data using independent components analysis. Journal of Neural Engineering 2006; 3(3): 208. https://doi.org/10.1088/1741-2560/3/3/003

[16] Milanesi M, Martini N, Vanello N, Positano V, Santarelli MF, Landini L. Independent component analysis applied to the removal of motion artifacts from electrocardiographic signals. Medical & Biological Engineering & Computing 2008; 46: 251-261. https://doi.org/10.1007/s11517-007-0293-8

[17] Vokorokos L, Ádám N, Madoš B. Non-Invasive Brain Imaging Technique for Playing Chess with Brain Computer Interface. International Journal of Computer and Information Technology 2014; 3(5).

[18] Xia Y, Cao X, Wu G-C, Cheng J. Acupuncture Therapy for Neurological Diseases: A Neurobiological View, Springer Science & Business Media 2010. https://doi.org/10.1007/978-3-642-10857-0

[19] Graimann B, Huggins JE, Levine SP, Pfurtscheller G. Visualization of significant ERD/ERS patterns in multichannel EEG and ECOG data. Clin Neurophysiol 2002; 113: 43-7. https://doi.org/10.1016/S1388-2457(01)00697-6

Received on 08-03-2017 Accepted on 04-04-2017 Published on 12-05-2017

https://doi.org/10.6000/1927-5129.2017.13.43