Spectral and Spatial Feature Extraction of Electroencephalographic (EEG) Data Using Independent Component Analysis (ICA)

Muhammad Jahangir$^{1,5}$, Syed Tanweer Iqbal$^2$, Rizwan-ur-Rehman$^{1,5}$, Nasir Ali Shah$^{3,5}$, Syed Mansoor Naqvi$^4$ and Imran Siddiqui$^{5,*}$

$^1$Govt. National College, No. 1, Karachi, Pakistan
$^2$NED University of Engineering and Technology, Karachi, Pakistan
$^3$Govt. College for Boys and Girls SRE Majeed, Stadium Road, Karachi, Pakistan
$^4$Aga Khan University Hospital, Karachi, Pakistan
$^5$Department of Physics, University of Karachi, Karachi-75270, Pakistan

Abstract: Purpose of this research is to extract features associated with human brain signal related to electroencephalographic measurements and classification of extracted EEG signals to the relevant brain region. EEG brain signals from 14 electrodes placed on the human scalp is recorded non-invasively using Emotiv EPOC / EPOC+. Scientific contextual EEG system with a sampling rate of 128 Hz. EEG data of human brain functions related to evoked motor imagery tasks consisting of two different classes of activities, namely imagination of right arm-movement i.e. arm down (termed here as PUSH) and arm up (termed here as PULL) for three healthy subjects is recorded. After pre-processing for noise and artifacts removal, the EEG signals associated with investigated evoked activities are extracted using Independent Component Analysis (ICA). The results obtained show good contrast plots for the extracted brain signals recorded on F7, FC5 and FC6 electrodes, decomposed on independent components, namely IC1, IC4, IC5, IC6. Classification of extracted features is mapped on to the motor imagery parts of human brain. The algorithm based on independent component analysis gives good results for feature extraction corresponding to evoked signals. Power spectra are also determined for the extracted independent components.

Keywords: Electroencephalogram, electrocorticogram, independent component analysis, brain computer interface, event related potential.

1. INTRODUCTION

Rapid technological advancement has resulted in data acquisition and analysis techniques of brain signals, such as electroencephalograms (EEGs) and electrocorticograms (ECoGs), has a profound impact on brain wave research, leading to various applications such as severe motor disabilities resulting from severed nerve connections to control prostheses for the movement of disabled body parts. This paradigm of scientific progress and technological advancement, this is, Brain-Computer Interface (BCI) provides an effective human control over devices such as computers, assistive appliances, neurophysiological disorders and sophisticated technology requiring fast and efficient human control. Although in its infancy, BCI has already started making differences in helping individuals in gaining their independency, an improvement in the quality of living.

Human brain is in a continuous state of dynamic working, generating signals associated with various activities which it can control and perform. Signals corresponding to different activities are in the form of weak brain electrical signals (microvolts) and are called electroencephalogram (EEG) signals. These brain activity measures (EE signals) are different from the measuring techniques used for measuring brain structure. Some of the brain structure analysis techniques are Magnetic Resonance Imaging (MRIs), Computerized Axial Tomography (CAT) scans and X-Rays. An EEG measuring device use surface sensors placed on subject scalp for measuring EEG signals or brain waves. The technique is non-invasive measuring EEG signals painlessly, i.e. without the need of opening the skull to implant sensors, as is done in invasive EEG recording.

Various brain activities and associated signals are classified according to their frequency band, viz., delta, theta, alpha and beta [1, 2]. The delta band of frequencies corresponds to frequency of 3 Hz or below. Signals lying in this band are highest in amplitude but lowest in frequency. It is a dominant rhythm in infants and/or in sleep state of adults. The theta frequency band corresponds to "slow activities" and is in the range 3.5 Hz to 7.5 Hz. Signals with these frequencies are normal in children up to the age of 13 years and in sleep state. Alpha frequency band lie in the range 7.5
Hz to 13 Hz. Beta frequency band correspond to fast activities with a frequency range of 14 Hz to 30 Hz. High-frequency brain activities are classified in the gamma frequency band lying in the range of 30 Hz to 80 Hz. There has been a considerable interest in studying the brain EEG signals corresponding to gamma frequency band. High frequency EEG signals associated with muscle activity (~20-300 Hz) completely mask the signals associated with the gamma frequency band. Therefore there is a growing appreciation in scientific community of the recorded EEG signals in the gamma frequency band being contaminated by neural signals associated with muscle activity.

EEG signals are used in a number of applications ranging from investigations of neurological disorders to Brain-Computer-Interfacing (BCI). EEG signals used for the diagnosis of Neurological disorders include pseudoseizures, epilepsy [3, 4]. Analysis of EEG data is also useful in the areas of research associated with the BCI [5, 6]. A number of research papers have been published, regarding awake and sleep states of a human brain [7], assessment of visual and audio alertness [8]. In the area of feature extraction from EEG data, a new method was proposed based on wavelet packet decomposition (WPD) [9]. Features associated with left and right hand movement were extracted from EEG data using wavelet transform and probabilistic neural network [10]. Monte Carlo simulation studies were also performed to remove the effect of forward model errors, different data sets of EEG, MEG and EEG/MEG are used for source localization [11]. The motor cortex, or M₁, area of the brain located in the rear portion of the frontal lobe, as shown in Figure 1a, is responsible for voluntary movements of the body. Prior to an actual movement, planning is done in the primary motor cortex area which communicates with other parts of the brain to assess the present state and position of the body. The skeletal muscles on the opposite side of the body are activated by neural impulses travelling across the body which are generated by primary motor cortex. This implies that the right and left sides of the body are controlled by the opposite sides of the human brain, viz. left and right hemispheres. All parts of the body are mapped to primary motor cortex area and are somatotopically arranged, i.e. there is a point-to-point correspondence of different parts of the body to specific points in primary cortex. In the motor cortex part of the brain, the amount of brain matter allocated to any part of the body depends upon the amount of control or sensitivity associated with that part of the body. An example of such correspondence can be seen in controlling complex movements such as those of the hands and fingers. In this case larger amount of cortical space is required as compared to that for trunk and legs whose motions are relatively simple (see Figure 1b).

2. MATERIALS AND METHODS

2.1. Experimental Setup and Pre-Processing of Recorded EEG Signals

In this study Emotiv EPOC / EPOC+: Scientific contextual EEG system is used to record EEG brain signals. The EEG signals from human brain are recorded by fixing electrodes on the subject scalp, 14
EEG electrodes are placed on the human scalp. An International 10-20 EEG electrode placement scheme is followed in this work [12-14] (see Figures 2 & 3). The actual experiment consists of recording EEG data at a sampling rate of 128 Hz and with an epoch time of ca. 1 minute (i.e. ca. 60 segments of EEG data are recorded in one minute duration), therefore a total of ca. 8000 data points are contained in data sets. To investigate the spectral region of our interest, the EEG data is filtered between 0.5 and 30 Hz. The line noise is suppressed by using a 50 Hz notch filter. The evoked activities belonging to motor imagery tasks corresponding to right or left arm down i.e. PUSH and up i.e. PULL are studied. EEG data of 3 normal subjects (all female) aging between 18-24 years is recorded using the experimental setup, schematically shown in Figure 4. The subject sits on a comfortable chair with arms in rest state. A visual cue is used to instruct the subject to perform left or right arm movement. The experiment consists of taking 10 trials for each subject separated by several minutes of relaxation. The recorded EEG data for different subjects performing same activity showed a considerable difference; therefore a training phase was conducted before actually taking the EEG data for signal analysis and decomposition. The subjects were requested to concentrate on the task given to them and at the same time maximum possible number of distractions were removed from the experimental area. Same cue description is followed while recording EEG data from different subjects. The scalp diagram showing 14 electrode positions AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 using 2 reference electrodes, (see Figure 3). The EEG data recorded simultaneously on all 14 electrodes constitutes a frame of EEG signals and in one second resulting in 128 such frames. The experimental setup is given in Figure 4. In addition to the EEG signals associated with neurological disorders and various other evoked activities, EEG signals associated with voluntary and involuntary activities e.g. muscular activities, blinking of eyes are some major sources of contamination of EEG data. The EEG data is also contaminated by power line electrical noises. These artefacts and sources have adverse affects on the useful features of the signals and before applying any signal processing technique, the recorded EEG data must be pre-processed for the elimination of unwanted signals. Several pre-processing steps are followed before any features are extracted, these include removal of artefacts, to reduce the dimensions of EEG data, signal averaging is performed, thresholds of the output are set, the resulting signal is amplified, and finally, edge detection. After pre-processing phase features are extracted using Independent Component Analysis (ICA).

Figure 2: The general system electrode placements. A 10–20 international system of electrode placement [11] is used.

Figure 3: In the experiment 14 electrodes are used whose positions are shown in the figure.

2.2. EEG Signal Processing and Feature Extraction

A feature is a unique or a characteristic measurement, structural component, transform extracted from a segment of a pattern [15]. Different feature extraction modalities are based on whether the features are extracted in space or time domain. The choice of a specific feature extraction scheme depends on the requirement that a certain feature or information is important for the classification of EEG data [16-18]. Some examples of such modalities are Principal Component Analysis (PCA), Independent Component Analysis (ICA), Autoregression technique (AR), Fast
Fourier Transform (FFT), Wavelet Transform (WT), Tensor Decomposition (TR). In this communiqué the authors have successfully extracted features from EEG data using Independent Component Analysis technique.

2.3. Independent Component Analysis

Independent Component Analysis (ICA) is a common statistical technique capable of linearly transforming any observed random data into underlying components such that they are made maximally independent of each other and at the same time addressing characteristics features in the distribution. The technique based on blind source separation problem and is used to extract features or independent components present in a given data set associated with some physical phenomenon.

Since both the source signals and their mixing procedure are both unknown so is termed as a blind source separation (BSS). ICA is a technique used for solving blind source separation problem and separates the independent components. The technique is based on the procedure in which the original data is transformed into a linear system such that the independent components present in the original system can be separated. The linear system developed is known as the un-mixing system. ICA is different from Principal Component Analysis (PCA) which is a correlation based transformation process. ICA is not only useful for its property of de-correlating various components but is also able to remove higher order statistical dependencies. The technique is outlined in the following paragraphs.

Given a set of random observations (EEG signals) measured in time domain or sample index, that is, \((x_1(t), x_2(t), \ldots, x_N(t))\) and are produced as a linear mixtures of independent components (independent features):

\[
\begin{align*}
\begin{bmatrix}
x_1(t) \\
x_2(t) \\
\vdots \\
x_N(t)
\end{bmatrix} &= 
\begin{bmatrix}
s_1(t) \\
s_2(t) \\
\vdots \\
s_N(t)
\end{bmatrix} \\
&= A 
\begin{bmatrix}
s_1(t) \\
s_2(t) \\
\vdots \\
s_N(t)
\end{bmatrix}
\end{align*}
\]  

(1)

where A is an unknown matrix whose elements are mixing coefficients

\[
A = 
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1N} \\
a_{21} & a_{22} & \cdots & a_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
a_{N1} & a_{N2} & \cdots & a_{NN}
\end{bmatrix}
\]  

(2)
If the components of $S(t)$ are such that at most one source is normally distributed then it is possible to extract the sources $S(t)$ from the received mixtures $x(t)$ [19, 20]. In ICA linear transformation matrix $W$ of the dependent sensor signals $x(t)$ is obtained which is the inverse of the unknown, mixing matrix $A$ such that the output is made as independent as possible

$$u(t) = Wx(t) = WAS(t)$$  \hspace{1cm} (3)

where $u(t)$ are estimate of sources and un-mixing signals, i.e.

$$u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_N(t) \end{pmatrix}$$  \hspace{1cm} (4)

The component independence is obtained using various mathematical techniques; in this work expectation value of kurtosis is calculated for the recorded and observed EEG data distribution.

### 3. RESULTS AND DISCUSSION

After pre-processing of the EEG time-series data, ICA technique is employed to separate various evoked features present in the EEG data using Fast ICA algorithm [19, 20]. In the present paper EEG data for ca. 60 trials or sub-epochs, each with 128 frames, are analysed. Recorded EEG signals pre-processing and application of Fast ICA algorithm are performed in Matlab workspace. Activity power spectrum associated with all 14 scalp electrodes AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 and for the decomposed independent components and topographic images are also generated in Matlab workspace. Plots of epochs vs. frames for the recorded EEG continuous data are also generated in Matlab workspace. Statistical analysis of the recorded EEG signal data for the scalp electrodes and for decomposed independent components is determined.

In this communiqué two ERP signals, i.e. brain signals associated with specific events or activities are analysed. These activities, associated with brain
signals, are for evoked potentials of arm down, here designated as PUSH and arm up, designated as PULL. EEG signals corresponding to the state of wakefulness, lying in 7-30 Hz rhythms associated with alpha and beta bands. In these state the signals are characterised by the presence of low voltage fast activity associated with visual, somatosensory and auditory cortices. EEG data is analysed for temporal and spatial separation by applying Fast Fourier Transform (FFT) and Independent Component Analysis (ICA) algorithm, respectively. The technique is applied on three EEG signals datasets recorded using three healthy subjects. In this study spectral and spatial features are extracted for all investigated datasets and topographic images showing spatial distribution of recorded signals are also developed. A topographic plot showing images of all 14 decomposed independent components of a dataset for evoked activity PUSH are shown in Figure 4. In Figure 4, a high unilateral contrast is seen at the position of electrode F7 decomposed on IC-5. Activity power spectrum and the identified spectral and spatial features from the analysis are shown in Figures 6-17.

The power spectrum for the electrode FC6 is shown in Figure 6 and is related to the evoked potential for PUSH activity for a subject performing the activity using left hand. A peak in the frequency range from 10-15 Hz can be seen in this figure. Application of ICA on the data set resulted in a similar line profile and an ERP appears on the independent component number 1 (IC-1) in the frequency range 10-15 Hz (see Figure 7). The power spectrum observed in Figure 8 is associated with electrode F7 which separates on IC5 as shown in Figure 9. In both figures the signals appear in the range 10-15 Hz but are broader as compared to those appearing in Figures 6 and 7. Similarly, the signals appearing in Figures 6 and 7 are for the subject who performed the said activity using right hand.

Figures 10 and 11 give the evoked activities associated with PUSH and are in the frequency range of 10-15 Hz. Signals are recorded on FC5 and separated on IC5. Brain signal on FC5 implies that activity was performed using right hand movement.
Figure 10: Activity power spectrum for electrode FC5 (PUSH-Dataset-3).

Figure 11: Activity power spectrum for independent component (IC-5) (PUSH-Dataset-3).

Figure 12: Activity power spectrum for electrode F7 (PULL-Dataset-1).

Figure 13: Activity power spectrum for independent component (IC-6) (PULL-Dataset-1).

Figure 14: Activity power spectrum for electrode F7 (PULL-Dataset-2).

Figure 15: Activity power spectrum for independent component (IC-4) (PULL-Dataset-2).

Figures 12, 14 and 16 show ERP signals for the evoked activity, PULL. The ERP signal appears on electrode F7 and separated on IC-6 or IC-4 as shown in Figures 13, 15 and 17. The analysis indicates that the signals are in the frequency range of 15-20 Hz and are for right hand arm up movement i.e. PULL activity. In Figure 8 the line profile is broad which can be explained as follows. The subject is queued to perform arm down movement i.e. PUSH activity. In the process the subject first generated a brain signal to raise right or left arm and then according to the queue generates a signal for arm down movement i.e. PUSH. This contributes to both the activities resulting in a broader line profile. Figure 18 shows the power spectrum for
PUSH activity recorded at electrode F7, decomposed on IC5. A high contrast is observed for the independent component 5. This suggests the decomposition of EEG signal recorded signal on IC5. Table 1 gives the statistical analysis for the recorded EEG data for different sets. For each data set statistical parameters are determined and Kolmogrove - Simronv tests are also performed to check the Gaussianity of the recorded data signals. Table 2 gives the statistical results for the decomposed independent components. The Kolmogrove - Simronv tests suggest that both recorded EEG signals and extracted features are non-Gaussian, this is indicated in Tables 1 & 2.

4. CONCLUSIONS

The objective of this study is to extract and classify EEG signals associated with event-related-Potentials (ERP) generated in primary motor cortex area for hand movement. EEG signals comprising six data sets for evoked activities PUSH and PULL are analysed using Independent Component Analysis (ICA) method. For the evoked activity PUSH the recorded brain signals lie in the frequency range of 10 – 15 Hz and for the activity PULL the recorded brain signals lie in the range 15 – 20 Hz. Good contrast plots are observed for the decomposed signals associated with electrodes F7, FC5 and FC6 which are decomposed on IC1, IC4, IC5, IC6. It can be concluded that the independent component analysis method gives good results for feature extraction corresponding to evoked signals. The study is beneficial in BCI domain where neural commands are transformed into control signals, these controlled signals then can be used in vast majority of applications ranging from mental and Physical disabilities to robotic control of machines.
Table 1: Descriptive analysis of the recorded EEG data sets. Six data sets are used in the investigation, three corresponding to PUSH and three to the PULL evoked activities. Column 2 shows the electrode number associated with the investigated activity

<table>
<thead>
<tr>
<th>Activity</th>
<th>Electrode</th>
<th>Data Points</th>
<th>Mean</th>
<th>Std-Dev</th>
<th>Var</th>
<th>Range</th>
<th>S</th>
<th>K</th>
<th>KS Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>PUSH</td>
<td>FC6</td>
<td>7936</td>
<td>0.0864</td>
<td>44.42</td>
<td>1974.00</td>
<td>448.50</td>
<td>-0.18</td>
<td>2.92</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>PUSH</td>
<td>F7</td>
<td>8064</td>
<td>0.0877</td>
<td>27.92</td>
<td>779.30</td>
<td>403.30</td>
<td>1.49</td>
<td>12.10</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>PUSH</td>
<td>FC5</td>
<td>8192</td>
<td>0.0718</td>
<td>21.99</td>
<td>483.50</td>
<td>517.90</td>
<td>1.33</td>
<td>19.90</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>PULL</td>
<td>F7</td>
<td>8320</td>
<td>0.0778</td>
<td>36.67</td>
<td>1344.00</td>
<td>541.40</td>
<td>0.90</td>
<td>5.98</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>PULL</td>
<td>F7</td>
<td>8064</td>
<td>-0.1420</td>
<td>32.08</td>
<td>1029.00</td>
<td>471.10</td>
<td>0.87</td>
<td>5.82</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>PULL</td>
<td>F7</td>
<td>8064</td>
<td>-0.0415</td>
<td>28.43</td>
<td>809.10</td>
<td>296.50</td>
<td>0.60</td>
<td>2.87</td>
<td>Distribution Non-Gussain</td>
</tr>
</tbody>
</table>

Table 2: Descriptive Analysis of the Extracted Features Corresponding to Various Independent Components

<table>
<thead>
<tr>
<th>ICs</th>
<th>Data Points</th>
<th>Mean</th>
<th>Std-Dev</th>
<th>Var</th>
<th>Range</th>
<th>S</th>
<th>K</th>
<th>KS Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC6-IC1</td>
<td>7936</td>
<td>0.0136</td>
<td>19.39</td>
<td>375.90</td>
<td>228.80</td>
<td>-0.31</td>
<td>1.79</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>F7-IC5</td>
<td>8064</td>
<td>-0.0006</td>
<td>3.97</td>
<td>15.80</td>
<td>52.32</td>
<td>0.13</td>
<td>1.82</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>FC5-IC5</td>
<td>8192</td>
<td>0.0036</td>
<td>4.98</td>
<td>24.82</td>
<td>149.70</td>
<td>2.75</td>
<td>62.10</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>F7-IC6</td>
<td>8320</td>
<td>-0.0059</td>
<td>4.52</td>
<td>20.46</td>
<td>53.14</td>
<td>-0.16</td>
<td>1.47</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>F7-IC4</td>
<td>8064</td>
<td>-0.0174</td>
<td>4.90</td>
<td>24.03</td>
<td>73.98</td>
<td>-0.15</td>
<td>3.79</td>
<td>Distribution Non-Gussain</td>
</tr>
<tr>
<td>F7-IC6</td>
<td>8064</td>
<td>0.0109</td>
<td>4.01</td>
<td>16.10</td>
<td>47.98</td>
<td>-0.24</td>
<td>2.68</td>
<td>Distribution Non-Gussain</td>
</tr>
</tbody>
</table>

ACKNOWLEDGEMENTS

The authors are thankful to department of Physics, University of Karachi for providing valuable assistance in arranging the EEG recording equipment. The authors are also thankful to Prof. Saud Wadood, Government National College, Karachi for frequent discussions on the subject and for moral and financial assistance.

REFERENCES

Spectral and Spatial Feature Extraction of Electroencephalographic
Journal of Basic & Applied Sciences, 2017, Volume 13 113


Received on 08-03-2017 Accepted on 23-03-2017 Published on 12-04-2017

https://doi.org/10.6000/1927-5129.2017.13.18

© 2017 Jahangir et al.; Licensee Lifescience Global.
This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.