Pre-processing and Feature Extraction Techniques for EEG-BCI Applications- A Review of Recent Research

Parthana Sarma 1, Prakash Tripathi2, Manash Pratim Sarma3, Kandarp Kumar Sarma4

1Gauhati University-Institute of Science and Technology, Gopinath Bordoloi Nagar, Guwahati - 781014, Assam. INDIA. parthana.sarma28@gmail.com
2Gauhati University-Institute of Science and Technology, Gopinath Bordoloi Nagar, Guwahati - 781014, Assam. INDIA. pcktripathi@gmail.com
3Department of Electronics and Communication Engineering, Gauhati University-Institute of Science and Technology, Gopinath Bordoloi Nagar, Guwahati - 781014, Assam. INDIA manashpelsc@gmail.com
4Department of Electronics and Communication Technology, Gauhati University, Gopinath Bordoloi Nagar, Guwahati - 781014, Assam. INDIA kandarpaks@gmail.com

Abstract: The electrical waveforms generated by brain named electroencephalogram (EEG) signals, require certain special processing for using them as part of applications. EEG signals need special pre-processing to enable brain computer interfaces (BCI) capture essential details of the signal and use them for specific applications, including deriving decisions. In this paper, we focus on some of the recent works reported in the area of EEG pre-processing. Further, we discuss some of the reported works related to feature extraction of EEG signals for application in drowsiness detection and development of assistive technologies for persons with special need.

Keywords: EEG signals, signal pre-processing, feature extraction, electroencephalography.

1. Introduction

The brain is the center of all activity in the human body the state of which can be monitored by recording the related signals. The brain signals resulting out of ionic flow variations within the neurons provide partial information about the physiological state of the human body and have been used for diagnostic purposes. Between the 17th and 23rd week of prenatal development, the neural activity of the human brain starts. The electrical signals generated by the brain indicate the effective functions of different related organs and hence is a promising way to communicate with the outside world using EEG signals [1]. This has led to the application of advanced digital signal processing techniques to electrical waveforms generated by brain named electroencephalogram (EEG) signals captured from a human subject. EEG records the electrical activity of the brain by capturing the voltage variations resulting out of internal neuron activity. Brain is completely covered by a network of different types of neurons, which are the processing unit of the data [2]. Each nerve cell consists of axons, dendrites and cell bodies. Nerve cells respond to stimuli and transmit information over long distances. An EEG signal is a measurement of currents that flow during synaptic excitations of the dendrites of many pyramidal neurons in the cerebral cortex. When brain cells (neurons) are activated, the synaptic currents are produced within the dendrites. This current generates a magnetic field measurable by electromyogram (EMG) machines and a secondary electrical field over the scalp measurable by EEG systems. The current in the brain is generated mostly by ionic activity of positive ions of sodium, Na+, potassium, K+, calcium, Ca++, and the negative ion of chlorine, Cl−. The direction of movement of these ions is governed by the membrane potential [2].

Certain diagnostic applications commonly stress upon the use of the spectral content of EEG. It covers a section of the neural oscillations contained in EEG signals. Most frequently, EEG is used to detect epilepsy, determined by noting the anomalous EEG readings. Disorders like sleep problems, coma, encephalopathies, and brain death are also related to EEG signal processing. Before the advent of high-resolution anatomical imaging techniques such as magnetic resonance imaging (MRI) and computed tomography (CT), EEG had been preferred as the first-line method of diagnosis for tumors, stroke and other focal brain disorders. Now the trend is toward greater use of anatomical imaging. Despite limited spatial resolution, EEG continues to be a valuable tool for research and diagnosis, especially when millisecond-range temporal resolution (not possible with CT or MRI) is required. EEG is an instantaneous and continuous indicator of brain's function, which has led to its extensive use in the field of biomedical signal processing [3]. For use of EEG, certain interface tools with the computer is required. In this context the BCI comes into consideration.

A Brain Computer Interface (BCI) is a communication system which is suitable for such a purpose. Here, messages or commands that an individual generate via brain, is transferred to a computer and subsequently related for use to control an external device. For example, by using EEG signals generated by the brain, a person can control the direction of movement of the wheel chair, the position of the cursor in the computer screen, etc. Success of BCI operation...
depends on the interaction of two adaptive controllers, user and system. The user must develop and maintain proper correlation between his or her intent and the signal features employed by the BCI; and the BCI must select and extract features that the user can control and must translate those features into device commands correctly and efficiently [3]. EEG based BCI has a number of applications in today’s world. Several assistive technologies have been developed based on EEG BCI. BCIs are mainly used for managing and regulating movements as seen in case of a motorized wheel chair or a prosthetic limb, restoring mobility in paralyzed limbs by electrically stimulating muscles, controlling home appliances, lights, television, room temperature, operate the door just by thinking, controlling a robotic car, play computer games, decoding brain activity to reproduce movements in robotic arms, controlling elements in virtual reality, walking in a virtual street by thought, typing a message on computer screen by concentrating on the display, controlling a computer cursor, for spelling words etc. [3]. This list is expanding rapidly, hence lots of research in this area are going on all over the world.

In this paper, we focus on some of the recent works reported in the area of EEG pre-processing. Further, we discuss some of the reported works related to feature extraction of EEG signals for application in drowsiness detection and development of assistive technologies for persons with special need.

The rest of the part is organized as follows. In Section 2, we include some basic theoretical concepts related to application development involving EEG signals. A detailed review of some of the important works related to EEG signal pre-processing is discussed in Section 3. In Section 4, some important works reported in the area of feature extraction of EEG signals are discussed. Next, in Section 5, a few likely scenarios of application development are discussed. The discussion is concluded in Section 6.

2. THEORETICAL CONSIDERATIONS

In this section, we discuss some of the related theoretical concepts.

2.1 EEG Signal

Figure 1 shows a few examples of the brain waves. EEG is typically a non-invasive (however invasive electrodes are often used in specific applications) method to record electrical activity of the brain along the scalp. Electrical recordings from the surface of the brain or even from the outer surface of the head demonstrate that there is continuous electrical activity in the brain [3]. Depending on the state of the brain, such as sleep, wakefulness or in any disease like epilepsy, the characteristics of EEG changes. Much of the time, the brain waves are irregular, and no specific pattern can be discerned in the EEG. The EEG signals are commonly decomposed into five EEG sub-bands: delta, theta, alpha, beta and gamma. Alpha waves are rhythmic and its frequency range is from 8 to 13 Hz. The amplitude of the alpha wave is low. Each region of the brain has the characteristic of alpha rhythm but mostly it is recorded from the occipital and parietal regions. It oscillates from adult in awake and relaxed state with eyes closed. Beta waves are irregular and its frequency range is greater than 13 Hz. The amplitude of the beta wave is very low. It is mostly recorded from temporal and frontal lobe. It oscillates from during the deep sleep, mental activity and is associated with remembering. Delta waves are rhythmic and its frequency range is from 8 to 13 Hz. The amplitude of the delta wave is high. It oscillates from the children in sleep state, drowsy adult and emotional distress occipital lobe. Theta waves are slow and its frequency range is less than 3.5 Hz. The amplitude of the theta wave is low-medium. It oscillates from adult and normal sleep rhythm. Gamma waves are the fastest brainwave frequency and its frequency range is from 31 to 100 with the smallest amplitude.

2.2 BCI

A Brain Computer Interface (BCI) is a communication system in which a person can use his or her generated EEG signal to control the external environment. BCI based applications provide a person with direct communication pathway between brain and external device in order to restore sight, hearing, movement, ability to communicate and cognitive function restoration.

The basic working block diagram of a BCI system is as shown in Figure 2.

Figure 1: EEG Sub-bands

Figure 2: BCI Block Diagram

A BCI system has an input (e.g. electrophysiological activity from the user), an output (i.e. device commands), components that translate input into output and a protocol that determines the onset, offset, and timing of operation.
Signals from the brain are acquired by electrodes on the scalp or in the head and processed to extract specific signal features (e.g., amplitudes of evoked potentials or sensory-motor cortex rhythms, firing rates of cortical neurons) that reflect the user’s intent. These features are translated into commands that operate a device (e.g., a simple word processing program, a wheelchair, or a neuroprosthesis).

2.3 Pre-processing and Feature Extraction

Pre-processing and feature extraction are two important steps in EEG signal processing. Pre-processing techniques help to remove unwanted artifacts from the EEG signal and hence improve the signal to noise ratio. A pre-processing block aids in improving the performance of the system by separating the noise from the actual signal. Subsequently, a feature extraction block helps to retrieve the most relevant features from the signal. These features will aid in the decision making mechanism in giving the desired output. A review of important works in this regard is included in the next section.

2.3.1 Least Mean Square (LMS) algorithm

Conventional filtering cannot be applied to eliminate the artifacts because EEG signal and artifacts have overlapping spectra. An adaptive process, in which the transfer function H(z) is adjusted according to an optimizing algorithm. The adaptation is directed by the error signal between the primary signal and the filter output. The most used optimizing criterion is the Least Mean Square algorithm. This algorithm is an application scheme widely used in practice due to its simplicity [25].

\[ W_{n+1} = W_n - \mu N e(n) \]  
(1)

2.3.2 Discrete Fourier Transform (DFT)

In time domain, representation of digital signals describes the signal amplitude versus the sampling time instant or the sample number. However, in some applications, signal frequency content is very useful otherwise than as digital signal samples. The representation of the digital signal in terms of its frequency component in a frequency domain, that is, the signal spectrum, needs to be developed.

The algorithm transforming the time domain signal samples to the frequency domain components is known as the discrete Fourier transform, or DFT. The DFT also establishes a relationship between the time domain representation and the frequency domain representation. Therefore, we can apply the DFT to perform frequency analysis of a time domain sequence.

In addition, the DFT is widely used in many other areas, including spectral analysis, acoustics, imaging/video, audio, instrumention, and communications systems [25].

\[ X(k) = \sum_{n=0}^{N-1} x(n)e^{-2\pi j nk/N}; 0 \leq n \leq N - 1 \]  
(2)

2.3.3 Discrete Cosine Transform (DCT)

The DCT is a transform that is very common when encoding video and audio tracks on computers. It has found applications in digital signal processing and particularly in transform coding systems for data compression/decompression. DCT takes correlated input data and concentrate its energy in just first few transform coefficients [25].

\[ F(u) = \sqrt{2/N} \sum_{i=0}^{N-1} A(i) \cos((\pi u / 2N)(2i + 1))f(i) \]  
(3)

3. Recent Works Related to Pre-processing of EEG Signals

Pre-processing of EEG signal is an essential and important step in any BCI based applications. It helps to remove unwanted artifact from the EEG signal and make it suitable for further processing.

In [1], the pre-processing technique used is Blind Source Separation (BSS). The EEG signal is initially filtered using a Notch Filter centered at 50 Hz, followed by BSS for EEG artifact removal. An FIR filter has been used between 8–25 Hz to get the required EEG signal for feature extraction. In [2], various pre-processing techniques for EEG signal have been reviewed. The first technique described is the use of basic filtering to remove unwanted artifact from the EEG signal. A basic Notch filter can be used to remove 50 Hz power supply signals. A band pass filter can also be used to get the desired band of frequency. The second technique discussed is the Adaptive Filtering. Here instead of a fixed filter, a filter that adapts to the spectrum of the recorded EEG is used for effective artifact removal. The last technique discussed is the BSS. In [3], basic filtering has been done to remove the unwanted artifacts from the EEG signal. The signal has been first high pass filtered with lower cut-off frequency 0.1Hz followed by a low pass filter with higher cut-off frequency 50 Hz. This band pass filtering aided in the removal of power line noise, EOG and EMG artifacts. In [4], band pass filtering has been applied to remove any signal that does not fall in the range of P300 frequencies of the EEG signal. Also data averaging over many trials has been done to improve the signal to noise ratio. In [5], the EEG signal has been filtered using a band filter between 8-12 Hz, which corresponds to the Mu rhythm frequency range.

Table 1 summarizes the pre-processing techniques that have been implemented in the aforesaid papers.

In [6], various pre-processing techniques have been reviewed. Techniques such as Weiner filters or Adaptive filters give better performance than conventional basic filtering of EEG signals. Weiner filters can be used for removal of EMG artifacts and Adaptive filtering can be used for removal of EOG artifacts and any background noise in general. Another effective technique discussed in the paper for artifact removal is the Independent Component Analysis (ICA), which is used for the removal of power line noise as well as EOG, EMG and ECG artifacts. In [7], a new method, multi-channel Weiner filter has been proposed for enhancement of EEG signal. It is an advanced method than the Blind Source Separation method. In [8], a new technique called Joint Approximate Diagonalization of Eigen-matrices method (JADE), has been implemented to calculate the independent components and hence remove unwanted artifacts from the EEG signal. In [9], an artifact removal technique has been proposed that incorporates Lifting
Wavelet Transform, rather than wavelet transform, with ICA technique for effective removal of artifacts from EEG signal. This method provides a better and efficient way to eliminate artifacts than traditional ICA method. In [10], the EEG signals was first band pass filtered to get the desired band of frequency, followed by ICA technique to remove any unwanted artifacts from the signal.

Table 2 gives a precise description about the various pre-processing techniques implemented in the papers discussed.

### I. Table 1: Summary of Pre-processing Techniques Implemented Using Conventional Filtering Methods

<table>
<thead>
<tr>
<th>Serial Number</th>
<th>Reference Number</th>
<th>Author</th>
<th>Year</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>[3]</td>
<td>J. P. Varghese</td>
<td>2009</td>
<td>Implemented basic band pass filtering for artifact removal and also extracted data epochs from the signal to study the event related EEG dynamics.</td>
</tr>
<tr>
<td>3</td>
<td>[4]</td>
<td>A. T. Campbell et al.</td>
<td>2010</td>
<td>Used basic filtering techniques to remove artifacts and also data averaging to improve SNR.</td>
</tr>
<tr>
<td>5</td>
<td>[5]</td>
<td>O. D. Eva et al.</td>
<td>2015</td>
<td>Basic band pass filtering has been implemented to get the desired Mu rhythm frequency range.</td>
</tr>
</tbody>
</table>

### II. Table 2: Summary of Pre-processing Techniques Implemented Using Adaptive Filtering

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Reference Number</th>
<th>Author</th>
<th>Year</th>
<th>Contribution</th>
</tr>
</thead>
</table>

In [11], an adaptive filter through wavelet transform has been implemented for artifact removal from EEG signal. The signal is first decomposed up to 8 levels using wavelet transform, following which it is subjected to adaptive filtering process. Wavelet reconstruction is applied to get the artifact free EEG signal. In [12], various pre-processing techniques have been discussed for effective artifact removal from EEG signals. Bipolar montage, Laplacian montage and Common Spatial Patterns are some of the techniques discussed in this paper. In [13], discrete wavelet transform has been implemented to remove noise from the EEG signal. The EEG signal is decomposed to 8 levels using discrete wavelet transform, which removes the noise and decomposes the EEG signal into its corresponding sub-bands. In [14], improved wavelet transform technique, called the wavelet packet transform, have been implemented. It is an improvement over the traditional wavelet transform because here while decomposing the signal into frequency sub-bands; it preserves the temporal structure of the signal. Also the samples of the sub-bands after wavelet packet representation, remains same as that of the original signal.

Table 3 gives a summary of the techniques described in the above mentioned papers.

### III. Table 3: Summary of Pre-processing Techniques Implemented Using Wavelet Based Approaches

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Reference Number</th>
<th>Author</th>
<th>Year</th>
<th>Contribution</th>
</tr>
</thead>
</table>
4. Recent works related to Feature extraction of EEG signals

Feature Extraction of EEG signal is an important step in any BCI based applications. It helps to extract the most relevant features from the EEG signal and thus giving a more precise description and hence making it suitable for further processing.

In [15], data set consisting of 416 trials of 500 ms length (sampling rate 1000 Hz) were taken. Only the first 10 DCT values were taken concentrating only on low frequency components hence there is no need of using any filters for high frequency noises. In [16], a technique based on EMD and DCT by exploiting temporal correlation that used EEG signals to detect the pre-stage of epileptic seizure using LS-SVM classifier is discussed. In [17], the last 25% of DCT coefficients are used. It is observed that the changing rate of frequencies is higher for Ictal signal than Interictal signal from Temporal.

Table 5 gives a precise description about the various feature extraction techniques implemented in the papers discussed. Table 4 gives a precise description about the various feature extraction techniques in the papers discussed. Singular Value Decomposition or SVD is an important technique of feature extraction used for EEG signals. The following few papers give an insight into how this technique has been implemented in a precise manner.

In [16], standard deviation features have been extracted from the larger and smaller singular values of SVD. The experimental results show that SVD with LS-SVM classifier outperform the state-of-the-art method based on EMD-based features in terms of accuracy, sensitivity, and specificity. In [17], they have reshape the EEG signals into a two dimensional matrix. The column vector of EEG signal is rearranged to square matrix for the computation of singular value components using SVD. From the singular values, they took non-zero diagonal values. In [18], a feature vector is formed by performing Higher Order Statistics (HOS) and complexity analysis on the signal. SVD is then used to reduce the dimension of the feature vector. A one-way ANOVA test was performed on the extracted feature vector to select statistically significant singular values. The selected singular values are used to train the Support vector machine (SVM) based classifier. In [19], the raw data contained reoccurring artifactual patterns with amplitudes significantly greater than the true EEG signals.

Table 5 gives a precise description about the various feature extraction techniques implemented in the papers discussed.

IV. Table 4: Summary of Feature Extraction Techniques Implemented Using DCT Techniques

<table>
<thead>
<tr>
<th>Sl. no.</th>
<th>Ref. no.</th>
<th>Author</th>
<th>Year</th>
<th>Contribution</th>
</tr>
</thead>
</table>

V. Table 5: Summary of Feature Extraction Techniques Implemented Using SVD Techniques

<table>
<thead>
<tr>
<th>Sl. no.</th>
<th>Ref. no.</th>
<th>Author</th>
<th>Year</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[16]</td>
<td>M. Z. Parvez et al</td>
<td>2014</td>
<td>Standard deviation features have been extracted from the larger and smaller singular values.</td>
</tr>
</tbody>
</table>

VI. Table 6: Summary of Feature Extraction Techniques Implemented Using DWT Approach

<table>
<thead>
<tr>
<th>Sl. no.</th>
<th>Ref. no.</th>
<th>Author</th>
<th>Year</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[23]</td>
<td>M. M. Shalker</td>
<td>2007</td>
<td>Filtered EEG signal offers a perfect success in the rejecting undesired frequencies</td>
</tr>
<tr>
<td>2</td>
<td>[21]</td>
<td>A. Nakate et al</td>
<td>2015</td>
<td>Decomposition of a signal into a number of scales</td>
</tr>
</tbody>
</table>

In [20], the input to this subsystem is the EEG signal data. Then by applying the DFT, the frequency components of that signal are obtained. Then the unwanted frequencies (i.e. the frequencies which are greater than 30 Hz) are removed. Then these clean frequency components are given as an input to the percentage power to calculate the feature vector. These feature vectors are given as an input to the classification stage. Paper [21] is devoted to the use of DWT both for signal preprocessing and signal segments feature extraction as an alternative to the commonly used DFT. In [22], emphasis is given to EEG signal segments and their analysis by a harmonic wavelet transform resulting in features standing for scales 1, 2 and 3 respectively covering three frequency bands. In [23], four levels of DWT using Debauches4 are implemented. Finally, to detect the frequency contained in each level, FFT is employed.

Table 6 gives a precise description about the various feature extraction techniques implemented in the papers discussed.

5. Application Development
In this section we discuss about two EEG based BCI systems.

5.1 Color based keyboard

Assistive BCI based technologies have helped people suffering from motor neuron diseases communicate with the outside world and have aided in their interactions with the external environment. One such assistive technology to be developed is a EEG based color based keyboard. Such a keyboard will help people suffering from diseases such as ALS to communicate with the world based on their EEG signals. To realize such a system the flow chart shown in Figure 3 has been proposed. The initial step is the data acquisition. Several data acquisition system is commercially available. Several Arduino boards are available for acquisition of biomedical signals. Arduino boards can take these biomedical signals in analog or digital form and can process these signals using the microcontroller on the board. Arduino programming language is used to program the microcontroller. These boards are advantageous because they offer a low cost solution. One of the economic EEG acquisition system available is the NeuroskyMindwave Mobile device. The basic system model can be seen as in figure 4. In [24] this device has been used to acquire the EEG signal. Following the data acquisition, based on the flow chart shown in figure 3, a pre-processing method has been implemented, which is based on the adaptive filtering technique. This technique is based on a basic LMS algorithm. Such an algorithm is used for its computational simplicity. The output thus obtained gives an artifact free EEG signal. This artifact free signal is fed to several feature extraction techniques. DCT, DFT, Central Moments, PCA, are some the techniques used to extract the most relevant features from the EEG signal. The features thus obtained give a precise description of the EEG signal which can be fed to an Artificial Neural Network for accurate decision making. The ANN will give the final classification of the different colors based on the EEG signal.

5.2 Drowsiness based Alert system

Another BCI application is to develop a system which can provide a certain pre warning or alert about an evolving drowsiness situation which is about to engulf a person engaged in execution of certain critical task. Such a system can prove to be effective in minimizing loss and damage of property and even life. It is proposed to develop a system which shall be capable of providing drowsiness alert by analyzing EEG signals. The system is constituted by an EEG acquisition system (shown in Figure 4), followed by pre-processing, artifacts removal, feature extraction and decision making. To realize such a system the flow chart shown in Figure 5 has been proposed.

6. Conclusion

In this paper various EEG pre-processing and feature extraction techniques have been discussed. Pre-processing and feature extraction are two most important steps in the development of any EEG-BCI system. Pre-processing techniques help in removal of unwanted artifacts from the EEG signal and feature extraction gives a precise and better description of the signal. The recent developments in this context have been elaborated in this paper. The paper also gives an amalgamated review of the various works that has been performed regarding pre-processing and feature extraction of EEG signal. These pre-processing and feature extraction techniques give a better representation and description of the EEG signal. The features thus obtained can be used in a decision making system, which essentially can be an Artificial Neural Network.

References


[17] M. Z. Parvez, M. Paul, “Classification of Ictal and Intertical EEG Signals”, School of Computer & Math, Charles Sturt University, Bathurst, SA, Australia


Author Profile:

Parthana Sarma, currently pursuing M-Tech in Signal Processing and VLSI from Gauhati University-Institute of Science and Technology, Guwahati, Assam, India, received her B.E in Applied Electronics and Instrumentation Engineering from Girijananda Chowdhury Institute of Management and Technology, Guwahati, Assam, India in 2013. Her area of interest includes Biomedical Signal Processing and VLSI design.

Prakash Tripathi, currently pursuing M-Tech in Signal Processing and VLSI from Gauhati University-Institute of Science and Technology, Guwahati, Assam, India, received his B-Tech in Electronics and Communication Engineering from Gauhati University-Institute of Science and Technology, Guwahati, Assam, India in 2014 and Diploma in Electronics and Communication Engineering from Assam Engineering Institute, Guwahati, Assam, India in 2011. His area of interest includes Signal Processing, Electronic Circuit design and Embedded System.

Manash Pratim Sarma, currently an Assistant Professor in Department of Electronics and Communication Engineering, Gauhati University-Institute of Science and Technology, Guwahati, Assam, India, received M. Sc. degree in Electronics from Gauhati University, Assam, India in 2008 and earned his M. Tech. in Electronic Design and Technology from Tezpur University, Assam, India in 2010. His areas of interest include wireless communication and VLSI system design.