EEG Feature Extraction using Parametric and Nonparametric Modelling Techniques

F. Shiman, S.H. Safavi, F.M. Vaneghi, M. Oladazimi, M.J. Safari, and F. Ibrahim, Member, IEEE

Abstract—A survey on most common methods for EEG feature extraction has been conducted in order to compare their performance. A unified method that makes the result comparable in the different research was not available. This paper discusses the best approaches for classifying the parametric and non-parametric methods. The finding indicates that parametric methods do not provide good performance for EEG signals and non-parametric methods may not provide detail information on EEG analysis as well as parametric methods.

I. INTRODUCTION

Exploring brain electrical activities using electroencephalogram (EEG) signals has increased recently. Brain-Computer Interface (BCI), classification of sleep stages, person authentication etc. are some applications of EEG signal analyzing. There are different methods by which features of EEG signals could be extracted and analyzed. These methods generally could be categorized in two basic categories, one of which is called “Non-Parametric Methods” and the other one is the “Parametric Methods”.

In non-parametric methods, which are the most common methods for analyzing EEG signals, Gaussian random procedure are detected by statistical possessions of EEG signals. Therefore, EEG signals can be explained through the first and second order moment. Amplitude Distribution, Interval Distribution, Correlation Analysis etc. are some instances of this approach.

In Parametric methods, a parametric model is applied for describing the signal which can result an enhanced estimators[1]. Autoregressive (AR), Moving average (MA) and Autoregressive moving average (ARMA), are some forms by which feature extraction of EEG signals is performed.

In this paper we evaluate 5 methods from extracted articles which big diversity of articles and methods could be found currently:

- Independent Component Analysis (ICA)
- Correlation Analysis
- Power Spectral Entropy
- Autoregressive Modeling
- ARMA Modeling

II. METHODS

It was decided to use different articles, extracting advantages and disadvantages of selected method. Finally, by categorizing the discussed approach into non-parametric and parametric methods, the general advantage and disadvantages of these two approaches is discussed. Hereunder based on the aforementioned perspective, the literature review of 5 different EEG feature extraction is expressed:

A. Independent Component Analysis (ICA)

ICA is a method capable of producing subcomponents of a multivariate signal presuming the sources of the signal to be mutually independent and also non-Gaussian. A mathematical formulation representing ICA is given as:

\[
x_1(t) = a_{11}s_1(t) + a_{12}s_2(t) + \cdots + a_{1n}s_n(t) \\
x_2(t) = a_{21}s_1(t) + a_{22}s_2(t) + \cdots + a_{2n}s_n(t) \\
x_k(t) = a_{k1}s_1(t) + a_{k2}s_2(t) + \cdots + a_{kn}s_n(t)
\]

Where \(x_i(t)\) indicates the \(k\) observations and \(n\) source signals are illustrated by \(s_l(t)\). The combination matrix consists of weight coefficients \(\{a_{ij}\}\) with size \(kn\) which are related to several undefined parameters (e.g. conduction of modeling volume providing the source localization using scalp electrodes).

ICA is used to estimate the source signals \(s_l(t)\) based on the recorded signals \(x_i(t)\), with assumptions that the source signals are non-Gaussian and statistically independent [2]. However, in some cases which the location of signal sources is needed, ICA is not effective [2]. Nevertheless, some solutions have been proposed to tackle this problem [2].

ICA has applications in many areas and was reported to give excellent performance in EEG analysis. An EEG data are recorded by electrodes which are electrical potentials in many special locations on the scalp and it is generally accepted that unidentified components of neural source activity are extracted from these scalp recordings with linear combinations [2].

B. Correlation Function

Investigating the correlation among different randomly selected variables can be achieved through a correlation function while the correlation is demonstrated as a function of temporal or spatial distance between two specific points.

Correlation functions of various random variables are sometimes identified as cross correlation functions to highlight that different variable are significant as they are created of cross correlations.
A correlation function can certainly referred to as one of the implications for a novel spectral analysis of EEG. Random data for correlation function expressed the significances of the data at the same time on the significance of the same data in autocorrelation analysis in terms of the common dependence. This equation can be described the cross correlation for \( x \) and \( y \) signals that indicated as

\[
\Phi_{xy}(\tau) = E\{x(t)y(t+\tau)\} \quad (4)
\]

Where \( \tau \) denotes the lag time. Providing the condition of \( x=\Phi_{xy}(\tau) \) represents the function of autocorrelation and also can be calculated for information of discrete values [1].

This method may be used for estimation of parametric model [3]. However, the estimation in this method is unbiased but not consistent[4]. Polarity coincidence correlation function [21], auto- or cross-averaging[5, 6] and complex demodulation [7] are some modifications which have been used for this method.

Correlation analysis could be applied for monitoring ischemic changes of EEG and researches show that the high precision can be discovered the ischemic event although its delay for computation of data [8]-[9]. Despite many advances which have been achieved in application of this method, a lot of neuroscientists still confirmed the cross-correlation between the performance of pairs of neural constructions to assume their functionality [10].

**C. Power spectral entropy (PSE)**

PSE measures the spectral complication of an uncertain system through information entropy. It assumes a random variable \( X \) as shape of the system for an uncertain system, and \( X \) describe as

\[
X = \{x_1, x_2, ..., x_n\} \quad (n \geq 1)
\]

The corresponding probability is

\[
P = \{p_1, p_2, ..., p_n\} 0 \leq p_i \leq 1, i = 1,2, ..., n
\]

Under Constraints

\[
\sum_{i=1}^{n} p_i = 1 \quad (5)
\]

Consequently, the information regarding entropy of the method can be represented as follows:

\[
H = -\sum_{i=1}^{n} p_i \ln p_i \quad (6)
\]

FFT transform converts the time-series of signals into the power spectrum referred to as information entropy of power spectrum as power spectral entropy [11].

PSE is described by a quantity of time indecision in frequency domain. The entropy value of power spectrum in EEG signals is small when the spectrum peak is narrow. It illustrates an apparent concussive rhythm in the signal that is resulted as small complication when wave is orderliness. While having smoother peak of the spectrum, its entropy is of greater value Thus, spectra structure of EEG signals can be revealed by PSE [11]. Many alterations applied in these methods to obtain consistent estimates [1].

EEG signals have been categorized into various frequency bands in terms of frequency contents. The brain dysfunction implies for having components of power and frequency within these specific bands. Regarding the spectral characteristics of EEG signal, time dependent variables of power spectra have been taken into consideration for further analysis of time variations[22]. One of the most significant benefits of this type of analysis is related to maintain all information contents of EEG when it transforms the artifacts with low frequencies into a narrow frequency spectrum [1].

In contrast, the disadvantages of these methods related to limitation of feature analysis in terms of labor intensiveness, inter-user variability, and storage problems. They also create linear and grayscale displays of spectral analysis to decrease the labor intensiveness [12], which can further define the sum of the presented data.

**D. Autoregressive (AR)**

AR modeling is one of the prominent parametric methods. It indicates that linear mixture of the past EEG samples plus an independent component (white noise) brings existing EEG sample.

The forward prediction of the EEG signal was accomplished using the following equation:

\[
x[n] = \sum_{i=1}^{p_1} a_i x[n-i] + e[n] \quad (7)
\]

Where \( e[n] \) described as prediction error (new information contained in the current EEG sample) [13].

Several features expressed the reason for reputation of AR modeling of EEG: i) Short-term EEG spectrum can be distinguished by AR process with sensible accuracy; ii) AR model is totally applied in time series analysis context; iii) Parameters of AR model are estimated by simple algorithms. As expected, AR models are suitable choice to analyze EEG for biomedical engineers [13].

Linear dependency of this model to past values, which is common in modeling approaches, may decrease accuracy of this method. However, Xuan Kong [13] increased the accuracy by adding a predictive part to above mentioned equation.

**E. ARMA Modeling**

An ARMA model is also one of the parametric models including an autoregressive (AR) part and a moving-average (MA) part. In addition it utilized as a predictor for time series feature values and also predicts at special time instance according to random distribution and past value(AR and MA parts) [14].

ARMA (m,n) model is:
\[ y_t = - \sum_{i=1}^{m} a_t^{(i)} y_{t-j} + \sum_{k=1}^{n} b_t^{(k)} e_{t-k} + e_t \quad (8) \]

\( y_t \) represents time sample of the EEG signal corresponding to a single channel, \( a_t^{(i)} \) is autoregressive and \( b_t^{(k)} \) is moving average parameter at discrete time instant \( t \) and \( m, n \) are the zeros and poles numbers respectively. White noise Gaussian is defined by \( e_t \) [14].

When frequency spectrum demonstrates both sharp peaks and deep nulls, this method is the most excellent model processes [14]. We can generate time-frequency spectra in parametric model with time varying parameters as the resolution is higher than short-time Fourier transform or Wavelet generated spectra. In majority of EEG applications ARMA generally referred to as a more common representation of the AR; regarding the fact that EEG signals seem to be correspond more to this model, it would be of sufficient accuracy applying such a model[15].

III. PERFORMANCE OF METHODS

For evaluating the performance of different EEG feature extraction methods, first of all, we should define the term “performance” and determine its indications. If we define method “performance” with indications as higher accuracy, precision and speed, then it is required to determine the application of method as well. In other word, higher speed could be considered as an indication of performance if the method is used in an EEG monitoring system, whereas higher accuracy is more crucial for BCI applications rather than EEG monitoring purposes. It is to say that, performance of different EEG feature extraction method could be measurable if specific application is desired.

As it was mentioned earlier, performing an extensive investigation regarding the five different methods with same application did not lead to any specific study. Some studies were found in which performance evaluation of specific EEG analyzing algorithm were conducted in comparison with MRI modality as a control/gold standard [16]. These studies were not enough to make a comparison table between five different methods.

However, the comparison could be performed in general, considering the theory and assumptions of each method, beside several experiments and study on these methods. In general terms, we can divide selected feature extraction methods into two categories of non-parametric and parametric. In consequence, Correlation Analysis and Power Spectral Entropy will be placed in first category and ICA, Autoregressive Modeling and ARMA Modeling will be considered in second one.

As it was mentioned before, ICA assumes EEG signals as a linear summation of several independent signal sources. This property makes ICA proper choice for person authentication and identification applications in which extracting source signals apart from EEG electrodes position. However, the number of electrodes must be assumed more than or equal to sources [2]. This means ICA is unable to identify the actual number of source signals. Thus, it is highlighted in blind signals.

The correlation function for random data describes the general dependence of the values of the data at one time on the values of the same data in the case of autocorrelation analysis (or of different data in the case of cross-correlation analysis) at another time [1]. However, in 1993, Westdorp conducted an investigation [17] showing that the EEG signal is normally correlated comparing with Laplacian approach. Nowadays, the attraction of this method is lost although there exist solutions to deal with this problem.

As it was mentioned earlier, AR is the most popular parametric method to analyze EEG signals. It can provide more details on spectrum data in comparison with non-parametric methods. Hosni, S.M. et al. showed in a study [18] that the best classification accuracy between AR, AR spectral analysis and power differences is AR model. This model is known as a proper one for clinical applications [15].

Disadvantage of analysis by spectral is due to need for long observation of the time to reach to the well spectral supposition and it may causes a confliction to the non-stationary treatment of brain EEG signals. Another main disadvantage is that in this method the desired end result is seldom due to some characteristic values that are mostly required, such as bandwidths, peak frequencies and fractional power quantities. Since these values are obtained from the power spectrum, the estimators may not be efficient, nor they are the statistical uncertainties. These drawbacks are essentially eliminated by applying parametric models. Parametric models indicate a considerable changes of the spectral properties in examples of placebo influenced EEG’s which were not detected by visual assessment of the EEG [19].

However, care must be taken that the defined models are descriptive and empirical, and consequently cannot reconstruct whole neurophysiologic specifications of the EEG. The defined model reproduces in particular linear character part of signal. This response is anticipated because the higher order coefficients are not considered. In addition using the model is not a proper choice when it comes to study amplitude distributions. Nevertheless, the model is quite adequate as long as spectral analysis is our main concern.

In general, linear analysis schemes, which were discussed in this paper, only utilize information retained in the autocorrelation function (i.e., the second-order cumulant). Additional information stored in higher-order cumulants is therefore ignored by assumption. Thus, while the power spectrum provides the energy distribution of a stationary process in the frequency domain, it cannot distinguish nonlinearly coupled frequency from spontaneously generated signals with the same resonance condition [24].

Table 1.shows the summary of advantages and disadvantages of above mentioned methods and their main applications to make it easier to compare the performances.
<table>
<thead>
<tr>
<th>Method Name</th>
<th>Type</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Main Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICA</td>
<td>Parametric</td>
<td>• Proper choice for person authentication and identification applications.</td>
<td>• Unable to identify the actual number of source signals as well as their locations.</td>
<td>• Person authentication and identification</td>
</tr>
<tr>
<td>Correlation Analysis</td>
<td>Nonparametric</td>
<td>• This method may be used for estimation of parametric model.</td>
<td>• EGG signal is fairly correlated normally in comparison with Laplacian analysis.</td>
<td>• Person authentication and identification</td>
</tr>
<tr>
<td>PSE</td>
<td>Nonparametric</td>
<td>• It has good usefulness in classification of functions and dysfunctions of EEG signals based on frequencies and power.</td>
<td>• It does not have good usefulness whenever the predictable pattern of EEG is desirable.</td>
<td>• Person authentication and identification</td>
</tr>
<tr>
<td>AR</td>
<td>Parametric</td>
<td>• Linearly dependency of this model to past values, which is common in modeling approaches, may decrease accuracy of this model.</td>
<td>• Distinguished the short-term EEG spectrum.</td>
<td>• Useful in real-time estimation</td>
</tr>
<tr>
<td>ARMA Modeling</td>
<td>Parametric</td>
<td>• The most efficient method for modeling courses having both sharp peaks and deep nulls within their frequency spectrum.</td>
<td>• It is inadequate for characterizing the coefficients of high orders being applied to EGG signal analysis.</td>
<td>• Useful in real-time estimation</td>
</tr>
</tbody>
</table>

IV. CONCLUSION

Two of the well-known ways for non-parameterized and parameterized method are the spectrum analysis and Autoregressive (AR) methods respectively[19]-[23]. Acclaim about the definite priority of methods according to their capability is very hard. The findings indicate each method has specific advantages and disadvantages which make it suitable for special application. Parametric methods, which assume a predefined pattern for EGG signals, may not provide high-quality performance for some EEG signals. In contrast, non-parametric methods, for instance, may not provide detail information on EEG analysis as much as parametric methods. It is crucial to make clear the application of the method, whenever the performance of analyzing method is discussed. Considering this, the optimum method for any application might be different.

ACKNOWLEDGMENT

This research is supported and funded by the Prime Minister’s Department, Malaysia, through a special grant Project No. 66-02-03-0061/H-00000-37039.

REFERENCES


A. Westdrop, "Volume conduction effects on correlation analysis of EEG data," 1993, pp. 150-152.


