

Epileptic Seizures Detection based on Empirical Mode Decomposition and Hilbert-Huang transform of EEG Signal

Mokhtar Mohammadi

Dept. of Computer Science
University of Human Development
Sulaimani, Iraq
mokhtar.mohammadi@uhd.edu.iq

Aso Mohammad Darwesh

Dept. of Computer Science
University of Human Development
Sulaimani, Iraq
aso.darwesh@uhd.edu.iq

Abstract— The electrical activities of brain fluctuate frequently and can be analyzed using electroencephalogram (EEG) signals. We present a new method for classification of ictal and seizure-free intracranial EEG recordings. The proposed method uses the application of multivariate empirical mode decomposition (MEMD) algorithm combines with the Hilbert transform as the Hilbert-Huang transform (HHT) and analyzing spectral energy of the intrinsic mode function of the signal. EMD uses the characteristics of signals to adaptively decompose them to several intrinsic mode functions (IMFs). Hilbert transforms (HTs) are then used to transform the IMFs into instantaneous frequencies (IFs), to obtain the signals time-frequency-energy distributions. Classification of the EEG signal that is epileptic seizure exists or not has been done using support vector machine. The algorithm was tested in 6 intracranial channels EEG records acquired in 9 patients with refractory epilepsy and validated by the Epilepsy Center of the University Hospital of Freiburg. The experimental results show that the proposed method efficiently detects the presence of epileptic seizure in EEG signals and also showed a reasonable accuracy in detection.

Keywords— Empirical Mode Decomposition; Hilbert-Huang transform; EEG Signal, Support Vector Machines

I. INTRODUCTION

Epilepsy is one of the most frequent brain dysfunction, it has a prevalence of 0.6% - 0.8% of the world population [1]. Epilepsy is characterized by a sudden, unexpected and transient electrical disturbance in the brain. The electroencephalogram (EEG) signal contains information which helps to predict epileptic seizures [2]. EEG signals are by nature non-stationary [3]. Therefore, for the correct feature extraction, several methods have been used in feature extraction task, most of them, based on measures taken from different time-frequency transforms, including, wavelet transform [4], [5], [6], Time-frequency distributions (Wigner Vile, short Time Fourier transform, Choi Williams) [7] and empirical mode decomposition [3]. In the latter a discussion is left open about features taken from the EMD, due to the decomposition is signal dependent, and then, the frequency band obtained in one component for

one signal may differs to another signal. In this manner, feature vector from a signal would not have relation to feature vector from another signal. Different approaches for feature extraction from EMD have been proposed in [8]. In recent years, the empirical mode decomposition (EMD) algorithm has become an established tool for the decomposition and time-frequency analysis of non-stationary signals [9]. Using EMD, the original signal is decomposed as a linear combination of intrinsic oscillatory modes, called Intrinsic Mode Functions (IMF)s. The IMFs are defined in such a way that the subsequent application of Hilbert transform, also known as Hilbert Huang transform (HHT), provides meaningful instantaneous frequency estimates [10]. The complete data driven nature of the algorithm makes it highly suitable for the non-stationary EEG signals. We here propose to use a combination of EMD and HHT to extract features from multichannel EEG signals. Standard EMD yields misaligned IMFs (corresponding to different frequency bands) for multiple channels and, hence, makes their comparison meaningless. This issue can be resolved, however, using the recently developed multivariate extensions of EMD [11] which processes the input signal directly in higher dimensional spaces where it resides. Combination of multivariate EMD (MEMD) and HHT used in [12] to extract spectral features from multichannel EEG signals. The method used in this work is the one proposed in [12] but with different features, may be used for epileptic seizure detection.

II. MATERIALS

The EEG database contains invasive EEG recordings of 21 patients (13M, 8F) suffering from medically intractable focal epilepsy. In 9 patients the source of epilepsy is located in the temporal lobe, other 6 suffer from frontal focal epilepsy and 1 of parietal epilepsy. The others 5 patients have two epileptic sources, 3 with temporooccipital epilepsy and 2 with fronto-temporal and temporo-parietal each. The data were recorded during invasive presurgical epilepsy monitoring at the Epilepsy Center of the University Hospital of Freiburg, Germany [13]. In order to obtain a high signal-to-noise ratio, fewer

artifacts, and to record directly from focal areas, intracranial grid-, strip-, and depth-electrodes were used. The EEG data were acquired using a Neurofile NT digital video EEG system with 128 channels, 256 Hz sampling rate, and a 16 bits A/D converter. Notch or band pass filters have not been applied in the acquisition stage. The available data include only 6 intracranial EEG channels (3 focal and 3 extra focal electrodes). In this study, the 6 EEG records were used. The records of 21 patients were supervised by experienced epileptologists and the onset and offset of the seizures are also available with the database.

III. METHODS

The proposed method includes a preprocessing stage, the MEMD computation, the feature extraction step, the feature selection and the classification phase.

A. Pre-processing

All EEG records were initially filtered, using code from EEGLab (Delorme and Makeig, 2004) to clean some artefacts [14]:

a 45-55Hz band-reject 12th-order filter to remove power line noise, a 60Hz cut-off low-pass 1st-order filter, and a 0.5Hz cut-off high-pass 5th-order filter to remove the dc component. Next, all EEG records were resampled to 128 Hz in order to reduce computation time of EMD decomposition. This operation does not have any influence on the results since the bandwidth of the signal of interest does not exceed the 60 Hz.

B. Empirical Mode Decomposition (EMD)

EMD is a general non-linear non-stationary signal decomposition method. The aim of the EMD is to decompose the signal into a sum of Intrinsic Mode Functions (IMFs). Given a signal $x(t)$, the algorithm of EMD is based on a sifting process that can be summarized as [12]:

1. Interpolate all the local maxima and in the signal with a cubic spline line, to produce the upper envelope.
2. Repeat for the local minima to produce the lower envelope.
3. Compute the mean of both envelopes m_1 .
4. Extract the detail.

$$h_1 = x(t) - m_1$$

5. Repeat the steps 1 to 4, and consider the detail h_i as the data, until detail h_1 can be considered an IMF.

6. After k iterations, the detail h_k is an IMF and is designated as: $IMF_1 = h_k$.

7. Iterate steps 1 to 6 on the residual r_j in order to obtain all the IMFs of the signal:

$$r_j = x(t) - IMF1 - IMF2 - \dots - IMFj$$

The procedure ends when the residual r_j is either a constant, a monotonic slope, or a function with only one extreme. The result of the EMD process produces N $IMFs$ ($c_1(t), \dots, c_N(t)$) and a residue sigma ($r_n(t)$)

$$x(t) = \sum_{n=1}^N c_n(t) + r_n(t) \quad (1)$$

In this work, each EEG records was divided in segments of 15 seconds, Then, $IMF1$ to $IMF4$ were calculated for each segment of every EEG records of each channel.

C. MULTIVARIATE EMPIRICAL MODE DECOMPOSITION

For multivariate signals, the local maxima and minima may not be defined directly. Moreover, the notion of ‘oscillatory modes’ defining an IMF is rather confusing for multivariate signals. To alleviate these problems, multivariate EMD (MEMD) was proposed in [12] which generates multiple n-dimensional envelopes by taking signal projections along different directions in n-dimensional spaces; these projections are then averaged to obtain the local mean. Once the mean signal is defined, the rest of the procedure is quite similar to the standard EMD. Then the proposed multivariate extension of EMD suitable for operating on general nonlinear and non stationary n-variate time series. In this work, $IMF1$ to $IMF4$ were calculated for 6 channel of every EEG record.

The resulting multivariate $IMFs$ obtained from the algorithm were matched so that the corresponding $IMFs$ belonged to the same frequency band. It should be noted that the analysis at the IMF level was only possible due to the matched IMFs obtained by MEMD; this cannot be achieved by applying univariate EMD on multiple channels separately. The lower order $IMFs$ capture fast oscillation modes of the signal, while the higher order $IMFs$ capture the slow oscillation modes.

D. FEATURE EXTRACTION

The purpose of feature extraction is to reduce the original data by measuring certain features that distinguish one input pattern from another. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features (also named feature vector). Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.[15]

Features extracted from EMD can be classified into two sets, measures taken directly from each *IMF* or measures taken from a transform of the *IMFs*, for instance Hilbert transform. The following parameters were calculated on each *IMF*:

1) Energy

Once the *IMF1, IMF2, IMF3* and *IMF4* were computed for all segments of each channel, the normalized energy of *IMFn* is calculated as:

$$E_n = \frac{\sum_{k=1}^K |IMF_n[k]|^2}{\sqrt{\sum_{n=1}^N \sum_{k=1}^K |IMF_n[k]|^2}} \tag{2}$$

2) Mean Frequency (MF) [13]

The Hilbert transform of a continuous *IMF_n(t)* is defined as:

$$Y_n(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{IMF_n(\tau)}{t - \tau} d\tau \tag{3}$$

Where P is the principal Cauchy value. In real applications, the Hilbert transform of a discrete signal *Y_n[k]* is calculated using the fast Fourier transform as proposed in [16]. Having obtained the Hilbert transform *Y_n[k]* of each *IMF_n[k]*, an analytic signal *Z_n[k]* can be expressed as:

$$Z_n[k] = IMF_n[k] + jY_n[k] \tag{4}$$

with instantaneous amplitude and phase defined as:

$$a_n[k] = \sqrt{IMF_n^2[k] + Y_n^2[k]} \tag{5}$$

$$\varphi_n[k] = \tan^{-1} \left(\frac{Y_n[k]}{IMF_n[k]} \right) \tag{6}$$

From equation 5, the instantaneous frequency can be calculated as:

$$\omega_n[k] = \frac{1}{2\pi} \frac{\varphi_n[k] - \varphi_n[k-1]}{T_s} \tag{7}$$

then the Weighted Average Instantaneous Frequency (*WAIF*) is obtained as follows [17]:

$$\langle WAIF \rangle_n = \frac{\sum_{k=1}^K a_n^2[k] \omega_n[k]}{\sum_{k=1}^K a_n^2[k]} \tag{8}$$

and the mean frequency of the original input signal can be obtained by the weighted contribution of the mean frequencies of the individual *IMFs*, as given in [18], by:

$$MF = \frac{\sum_{j=1}^K \|a_j\| (WAIF)_j}{\sum_{j=1}^K \|a_j\|} \tag{9}$$

In this stage the *MF* of input signal was calculated as a feature.

E. FEATURE SELECTION

In order to reduce the dimensionality problem, the median of the individual values of each series for the six channels were initially computed. This measure of central tendency was selected since it is robust to outlier values. Thus, the number of the total features series is reduced to 5.

F. CLASSIFICATION

Support Vector Machines (SVMs) is a supervised learning method which performs classification by constructing an N-dimensional hyper plane that optimally separates the data into two classes [19]. The SVM approach is computationally efficient in training and classification, and it comes with a learning theory that can guide real world applications. The machine is presented with a set of training examples, (x_i, y_i) where the x_i is the real world data instances and the y_i are the labels indicating which class the instance belongs to. For the two class pattern recognition problem, $y_i = +1$ or $y_i = -1$. A training example (x_i, y_i) is called positive if $y_i = +1$ and negative otherwise. SVMs construct a hyper plane that separates two classes and tries to achieve maximum separation between the classes. Separating the classes with a large margin minimizes a bound on the expected generalization error. The epilepsy detection is carried out using LibSVM and the tool has been developed using MATLAB. LIBSVM is simple, easy-to-use and efficient software for classification and regression [20]. We tried the following procedure in this stage:

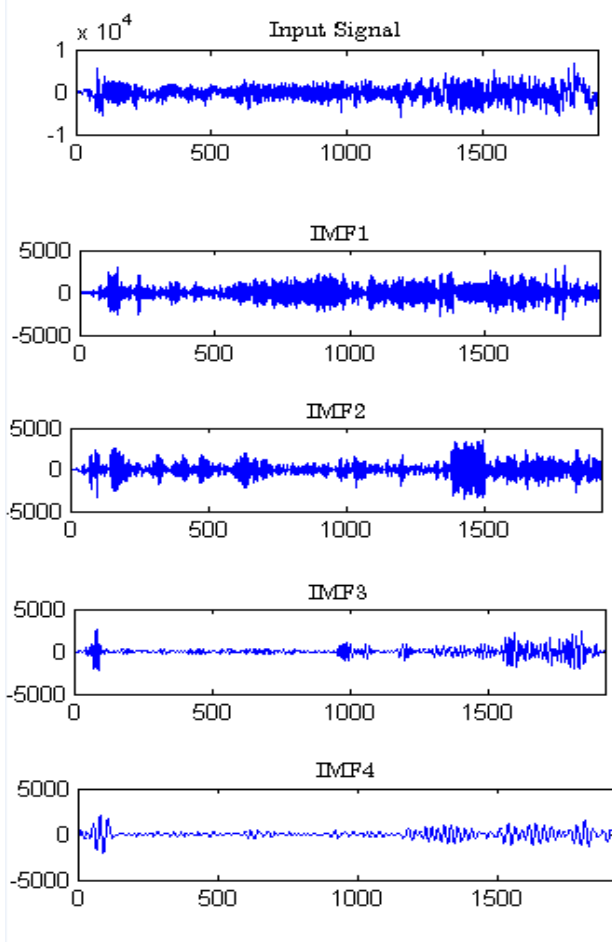


Fig. 1. EEG signal with an epileptic seizure corresponding to patient 4, channel 1, IMF1 to IMF4

1. Transform data to the format of an SVM package.
2. Conduct simple scaling on the data.
3. Consider the radial basis function (RBF) kernel $K(x, y) = e^{-\gamma \|x - y\|^2}$
4. Use cross-validation to find the best parameter C and γ . C is the penalty parameter of the error term.
5. Use the best parameter C and γ to train the whole training set.
6. Test.

It is not known beforehand which C and γ are best for a given problem; consequently some kind of model selection (parameter search) must be done. The goal is to identify good (C, γ) so that the classifier can accurately predict unknown data (i.e. testing data). The prediction accuracy obtained from the "unknown" set more precisely reflects the performance on classifying an independent data set. An improved version of this procedure is known as cross-validation. We tried a "grid-

search" on C and γ using cross-validation. Various pairs of (C, γ) values are tried and the one with the best cross-validation accuracy is picked.

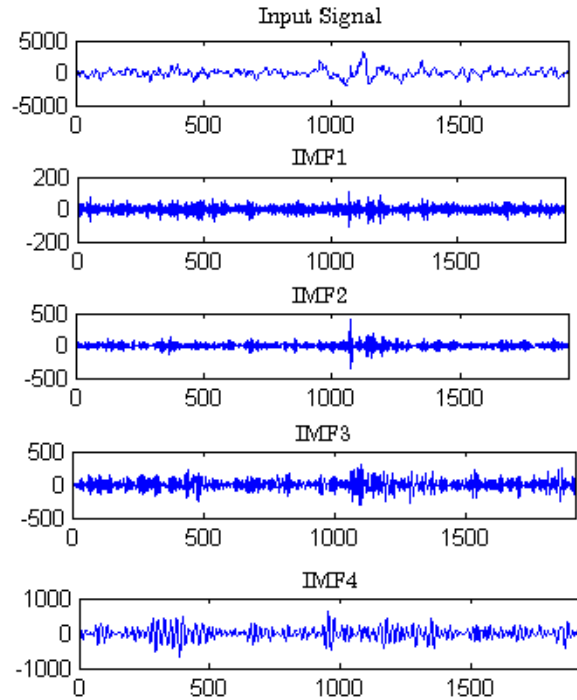


Fig. 2. Normal EEG signal corresponding to patient 4, channel 1 IMF1 to IMF4

IV. EXPERIMENTAL RESULTS AND ANALYSIS

We performed the time-frequency analysis of the EEG signals using the combination of multivariate EMD and HHT to obtain the mean frequency and the Energy of IMF1 to IMF4 of the signal. The Energy of each IMFs is computed by using the equation (1). In each IMFs the energy of the seizure signal is much higher than the energy of the non-seizure signal. Mean Frequency was a good feature to discriminate normal EEG signals with higher energy from seizure EEG signals. Average value of the feature vector for each patient are shown in Table I, normal state and Table II Epileptic state. In this method 60 percent of the available signal is used for training phase and the remaining 40 percent is for testing phase. The detection accuracy of this method is given in table III. The Accuracy of the SVM is defined as the ratio of correctly detected signals to the total number of signals.

Table 1: extracted features from ictal data

Patient	Mean Fr	En-IMF1	En-IMF2	En-IMF3	En-IMF4
1	0.10	1188	2893	5644	7620
2	0.07	966	4418	13356	23826
3	0.10	3346	15377	22173	24554
4	0.09	8544	37892	75698	123127
5	0.10	970	761	1398	2160
6	0.095	1533	4495	9175	16049
7	0.9	1839	8443	11807	12124
8	0.85	1341	5037	8164	12718
9	0.10	1752	6380	9480	8667

Table II: extracted features from non-ictal data

Patient	Mean Fr	En-IMF1	En-IMF2	En-IMF3	En-IMF4
1	0.11	1300	2964	6497	6491
2	0.09	185	637	2035	5170
3	0.09	481	1625	3861	8368
4	0.08	354	1449	3912	10390
5	0.09	207	367	613	1539
6	0.091	451	1190	2537	6330
7	0.08	242	936	5322	6263
8	0.94	1358	5368	7433	5071
9	0.11	740	2413	4681	4208

Table III: CLASSIFICATION ACCURACY OF EEG SIGNAL USING SVM

Patient	Accuracy
1	97.46
2	97.56
3	100
4	100
5	99.73
6	94.11
8	97.46
9	97.46

V. CONCLUSIONS

We have used a combination of MEMD and HHT to calculate the mean frequency and Energy, as a spectral features, in order to identify the seizure condition in EEG signals. The matched IMFs facilitate the comparison between them, which was not be possible with standard univariate EMD alone. RBF kernel function is used in the SVM classifier to classify/detect seizure EEG signal and normal EEG signal. Accuracy of the classifier is computed. The accuracy of this project is much better than other results available in the literature based on EMD.

VI. ACKNOWLEDGMENT

The authors are very grateful to the Epilepsy center of the

University Hospital of Freiburg, Germany, for their consent to use the invasive EEG recordings in this work.

VII. REFERENCES

- [1] C.G. Cassandras and S. Lafortune. Introduction to discrete event systems, Second Edition.Springer, 2007 .
- [2] Ouali M. S., Ait-Kadi D., and Rezg N. Fault diagnosis model based on Petri net with fuzzy colors. Computers and Industrial Engineering , 37(12):173176, 1999.
- [3] Y. Power and P. A. Bahri. A two-step supervisory fault diagnosis framework. Computers and Chemical Engineering, 28(11):21312140, 200.
- [4] Lefebvre D., Delherm C., Fault detection and isolation of discrete event systems with Petri net models, IEEE TASE, Vol. 4, nu. 1, pp. 114118, January 2007.
- [5] Lefebvre D., Delherm C., Diagnosis of DES with Petri nets models, accepted for publication IEEE TASE , 2006.
- [6] C.A. Petri. Kommunikation mit Automaten. PhD thesis, Institut fur Instrumentelle Mathematik, Schriften des IIM, No. 3, Bonn, Germany, 1962. [cited at p. 4, 24]
- [7] Bello, M. G. "Enhanced training algorithms, and integrated training/architecture selection for multi layer perceptron networks," IEEE Trans.on Neural Net., vol. 3, pp. 864-875, 1992.
- [8] Rumelhart, D. E., Hinton, G. E. and Williams, R. J, "Learning internal representations by error propagation," In Parallel Distributed Processing, Cambridge, MA: MIT Press, vol 1, pp. 318-362.
- [9] Werbos, P. J. "Back-propagation: Past and future," Proceeding of International Conference on Neural Networks, San Diego, CA, 1, pp.343-354, 1988.
- [10] Salvetti, A. and Wilamowski, B. M., "Introducing Stochastic Process within the backpropagation Algorithm for Improved Convergence", presented at ANNIE'94 - Artificial Neural Networks in Engineering, st. Louis, Missouri, USA, November 13-16, 1994; also in Intelligent Engineering Systems Through Artificial Neural Networks vol 4, pp. 205- 209, ed. C. H. Dagli, B. R. Fernandez, J. Gosh, R.T. S. Kumara, ASME PRESS, New York 1994.
- [11] Hagan, M. T. and Menhaj, M., "Training feedforward networks with the Marquardt algorithm", IEEE Transactions on Neural Networks, vol.5, no. 6, pp. 989- 993, 1994.
- [12] S. Ould El Mehdi, E. Leclercq and D. Lefebvre, Petri nets design and identification for the diagnosis of discrete event systems, IAR-ACD Annual Meeting Nancy, workshop on Advanced Control and Diagnosis, 2006.