Features Extraction and Classification of EEG Signals Using Empirical Mode Decomposition and Support Vector Machine

Noran M. El-Kafrawy, Doaa Hegazy, and Mohamed Fahmy Tolba

Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt
nel-kafrawy@acm.org, doaa_hegazy@fcis.asu.edu.eg, fahmytolba@gmail.com

Abstract. Interpreting brain waves can be so important and useful in many ways. Having more control on your devices, helping disabled people, or just getting personalized systems that depend on your mood are only some examples of what it can be used for. An important issue in designing a brain-computer interface (BCI) is interpreting the signals. There are many different mental tasks to be considered. In this paper we focus on interpreting left, right, foot and tongue imagery tasks. We use Empirical Mode Decomposition (EMD) for feature extraction and Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel for classification. We evaluate our system on the dataset 2a from BCI competition IV, and very promising classification accuracy that reached 100% is obtained.

1 Introduction

Consider the potential to manipulate computers or machinery with nothing more than thoughts. Our brains are filled with neurons. Every time we think, move, feel or remember something, our neurons are at work. That work is carried out by small electric signals that zip from neuron to neuron. Although the paths the signals take are insulated, some of the electric signal escapes. Scientists can detect those signals, interpret what they mean and use them to direct a device of some kind [1]. There are several applications that could make use of these data. Such as: assistive technology, virtual reality [2], game controlling [3] and robotics [4, 5]. We are interested in understanding these waves to be able to use them as an input to other systems. Several studies have been made in this field with varying results. We will focus on features extraction and classification of the brain waves.

In this paper we propose a method for classification of the EEG signals. This method consists of three steps: (1) Pre-Processing for eliminating unwanted noise. (2) Features Extraction using Empirical Mode Decomposition (EMD) [6]. (3) Classification using Support Vector Machine (SVM) [7]. The method is applied to real human data, and better results relative to the state of art algorithms are obtained.

In section two, we give related work and the available techniques used to solve the problem of feature extraction and recognition of EEG signals. Details of our

A.E. Hassanien et al. (Eds.): AMLTA 2014, CCIS 488, pp. 189–198, 2014.

proposed model are presented in section three. The data set used, numerical experiments and results are shown in section four, where we show the results obtained by the selected method. Finally, conclusions are driven and presented in section five.

2 Related Work

Many researchers use different techniques in feature extraction [8, 9, 10, 11]. Toka Fatehi et. al [12] gave an overview of different methods extracting features from an EEG signal. These methods include: Time Analysis, Frequency Analysis, Time-Frequency Analysis and Time-Frequency-Space Analysis.

In [12], it is shown that the best classification results were obtained from the space-time-frequency analysis.

However, feature extraction methods aren't restricted on Fourier Transform presented in [12]. In [13], Hualou Liang et. al. showed that the Fourier-based methods are designed for the frequency analysis of stationary time series, while in neurobiology the researchers deal with time series data that are non-stationary. Thus, Fourier-based methods can have limited use in revealing the underlying neurophysiological variations in such data [13]. Moreover, it is mentioned in [13] that the major drawback of Fourier-based approaches is that the basis functions are fixed, and therefore cannot capture any time-varying characteristic of neural signals.

In [14], Wang et. al. addressed the problem of analyzing EEG signals and classifying them into different classes of different motor imaginary tasks. They used Hilbert-Huang transformation method in feature extraction: It consists of two main steps (1) EMD. (2) The Hilbert spectral analysis. They used BP neural network in classification. They classified three different tasks (left hand, right hand and foot) and their best classification rate was 93.8%.

Demir, B. et. al. addressed in [15] a similar problem. They analyzed images using EMD (the same method for feature extraction used in [14]). Their best classification rate was 100%. However, our proposed method mostly differ from this mentioned in [14] in the classification part. We use the SVM with radial basis function (RBF) kernel to classify the feature vector obtained from the EMD process.

In the model proposed in this paper, EMD is used for features extraction. EMD was first introduced by Huang et. al. [6] as a signal decomposition method. EMD provides an alternative to traditional time-frequency methods and its main function is to decompose a signal into a collection of oscillatory modes, called intrinsic mode functions (IMFs). As mentioned in [16], EMD has a useful feature that it relies on a fully data-driven mechanism that does not require any prior known basis. By this feature it differs from the traditional signal analysis tools, such as Fourier or wavelet-based methods, which require some predefined basis functions to represent a signal.

Classification of EGG signal is an emerging and important field in signal processing. Different methods for classifying are proposed and used in the literature. In [17], Lotte et. al. presented a review on classification algorithms for

EEG-based brain-computer interfaces. Moreover, in [17], a survey of the classification algorithms used to design BCI systems is presented. They are divided into five different categories: Linear Classifiers, Neural Networks, Nonlinear Bayesian Classifiers, Nearest Neighbor Classifiers and Combination of Classifiers.

They also provid some guidelines to choose a classifier. Several measures of performance have been proposed in BCI, such as accuracy of classification, Kappa coefficient, Mutual Information, sensitivity and specificity. The most common one is the accuracy of classification and the percentage of correctly classified feature vectors, and this is the only considered measure.

Synchronous BCI is the most widely spread. In [17], three kinds of algorithms proved to be efficient in this context, namely, SVM, dynamic classifiers and combination of classifiers. SVM is very efficient regardless of the number of classes; this is because it has good properties like: regularization, simplicity and immunity to the curse-of-dimensionality. It is said to be the most appropriate classifier to deal with feature vectors of high dimensionality (e.g.: large number of time segments).

3 Proposed Work

Our system starts by reading a previously saved database of EEG signals. These signals have four different motor imaginary tasks (left hand, right hand, both feet and tongue). Our job is to classify these signals into four different classes. We apply EMD to extract features from these signals (IMFs), then we use the SVM classifier with the RBF kernel. Figure 1 shows the main steps of the proposed EEG signal classification model.

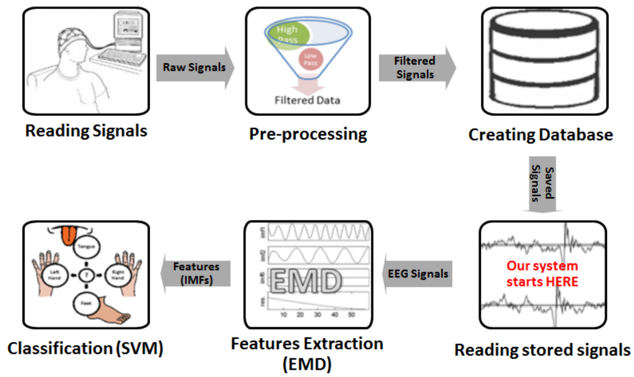


Fig. 1. Block Diagram of the Work

3.1 Pre-Processing

The signals are sampled with 250 Hz and bandpass-filtered between 0.5 Hz and 100 Hz. The sensitivity of the amplifier was set to 100 μV. An additional 50 Hz notch filter was enabled to suppress line noise.

3.2 Feature Extraction: Empirical Mode Decomposition

A Single Channel Case. EMD is a method of breaking down a signal without leaving the time domain. It is useful for analyzing natural signals (like brain waves – EEG). It decomposes the signal into intrinsic mode functions (IMFs), which is easier to analyze.
An IMF is a function with the following properties:

1. Has only one extreme between zero crossings, and
2. Has a mean value of zero.

The EMD uses a process called sifting to decompose the signal into IMFs. The following steps [18] explain its procedure. For a signal $x(t)$, let m_1 be the mean of its upper and lower envelopes as determined from a cubic-spline interpolation of local maxima and minima as shown in Figure 2.

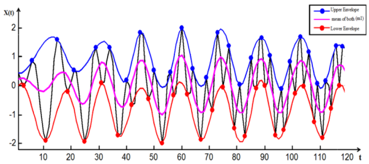


Fig. 2. Iteration 0: Upper & lower envelopes with their mean (m_1)

The difference between the data and m_1 is the first component h_1.

$$h_1 = x(t) - m_1 \tag{1}$$

The sifting process is repeated with h_1 treated as the main data. m_{11} is the mean of the upper and lower envelopes of h_1.

$$h_{11} = h_1 - m_{11} \tag{2}$$

This sifting process is repeated k times, until h_{1k} is itself an IMF.

$$h_{1(k-1)} - m_{1k} = h_{1k} \tag{3}$$

$$c_1 = h_{1k} \tag{4}$$

c_1 contains the shortest period component of the signal. It is separated from the rest of the data:

$$x(t) – c_1 = r_1 \tag{5}$$

And this process is repeated on $r_j : r_1 - c_2 = r_2, \cdots, r_{n-1} - c_n = r_n$ until we get all the possible IMFs of the signal as shown in Figure 5.

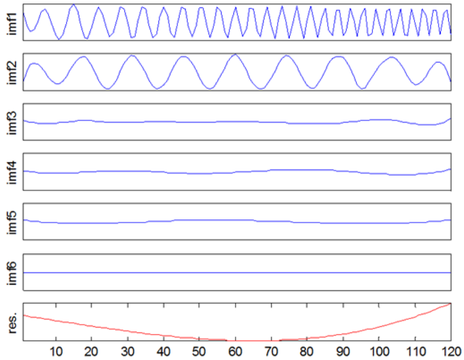
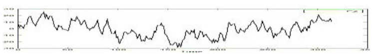


Fig. 3. IMFs obtained using EMD

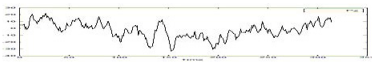
Characteristic features are obtained from the first two IMFs. As they contain the most features (high frequencies).These are used as input feature vectors for the classifier.

Multichannel Case. To analyze multichannel EEG signals recorded in our experiment synchronously, we decompose each channel separately to prevent any possible oscillatory information leaking among the channels[19]. We then calculate the 1st two IMFs of each channel, sum them and use the sum as a feature vector for the SVM classifier.

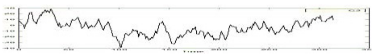
Figure 4 represents an event recorded from five channels. Each of them will be separately analyzed.



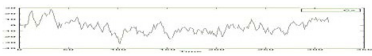
(a) Fz Signal



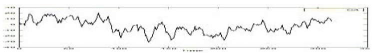
(b) Pz Signal



(c) C3 Signal



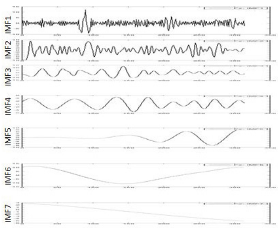
(d) Cz Signal



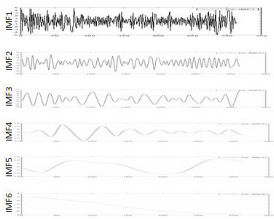
(e) C4 Signal

Fig. 4. Signals obtained from Fz, Pz, C3, Cz, C4 nodes correspondingly

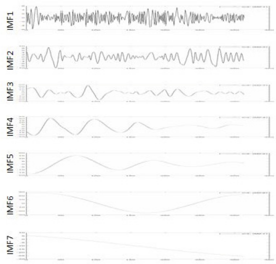
Figure 5 shows the IMFs obtained after applying EMD algorithm on the signals shown in Figure 4.



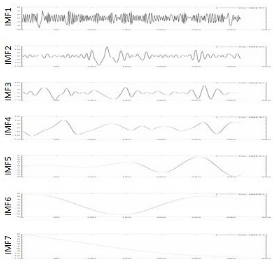
(a) IMFs from Fz Signal



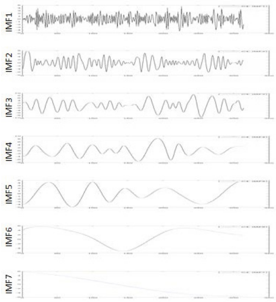
(b) IMFs from Pz Signal



(c) IMFs from C3 Signal



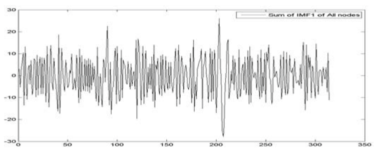
(d) IMFs from Cz Signal



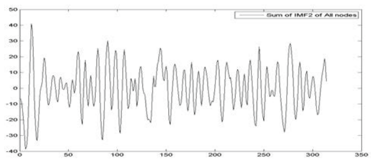
(e) IMFs from C4 Signal

Fig. 5. IMFs obtained from each signal after applying the EMD algorithm

Figure 6 shows the summation of the 1st two IMFs of each node, which contains the highest frequencies and used as a feature vector for the SVM classifier.



(a) The Sum of the 1st IMF



(b) The Sum of the 2nd IMF

Fig. 6. The summation of the 1st two IMFs of each node, which are used as a feature vector for the classifier

3.3 Classification: Support Vector Machine

Due to the use of a large number of time segments, our feature vectors are of very high dimensionality. SVM is the most appropriate classifier to deal with such feature vectors [17]. There are several kernel functions that could be used with SVM: Linear, Quadratic, Radial Basis Function (RBF), Polynomial and Multi-layer Preceptron (MLP). As mentioned in [18], the Guassian or Radial

Basis Function (RBF) kernel are the most used in BCI classification tasks. The corresponding SVM is known as Gaussian SVM or RBF SVM and is given as:

$$k(x, y) = \exp\left(\frac{-\|x-y\|^2}{2\sigma^2}\right) \tag{6}$$

In(6) x & y can be recognized as the squared Euclidean distance between the two feature vectors, while σ is a free parameter.

4 Results and Discussion

4.1 Data Set

The publically online available BCI competition 2008 dataset [20] is used in this work. The data set consists of 9 subjects. These subjects has four different motor imaginary tasks: the imagination of left hand movement (class 1), right hand (class 2), both feet (class 3), and tongue (class 4). Each subject recorded two sessions on two different days. Each session consists of 6 runs separated by short breaks. One run consists of 48 trials (12 for each of the four possible classes), This yields to a total of 288 trials per session. Twenty-two electrodes were used to record the EEG signals; the montage of these electrodes is corresponding to the international 10-20 system. As stated in [20], the subjects were sitting in a comfortable armchair in front of the screen of the computer. At the beginning of a trial ($t = 0s$), a fixation cross was shown on the black screen and a short acoustic warning tone was presented. After two seconds ($t = 2s$), a cue in the form of an arrow pointing either to the left, right, down or up (corresponding to one of the four classes left hand, right hand, foot or tongue) appeared and stayed on the screen for $1.255s$. This prompted the subjects to perform the desired motor imagery task. There was no feedback provided. The subjects were asked to carry out the motor imagery task until the fixation cross disappeared from the screen at $t = 6s$. A short break followed where the screen was black again.

4.2 Experimental Evaluation

In our experiment we used 80% of the total number of feature vectors as a training sample, while the remaining 20% as a test sample. Given a total number of features equal to fifty-four [1], the training sample is of size forty-three feature vectors, while the test sample is eleven. The feature vector in our problem is of a high dimensionality (three-hundred-fourteen) due to the use of a large number of time segments.

We applied EMD in order to get the required features (IMFs). We used the first and the second IMFs as feature vectors: as they contain the highest frequencies and thus the most features. Ten-fold Cross validation was first applied

[1] The number of training and test samples as well as feature size are experimentaly evaluated. However, not shown here due to the limited space.

to select the best parameters to train the SVM classifier. These parameters were chosen after several searches, this is to insure a global minimum (some of the results for the parameters are shown in Table 1). The default parameters [1,1] are close to the optimal for this data and partition. We used these optimal values in training the new SVM classifier and obtaining our final results. SVM classifier was then applied on the feature vectors using these optimal parameters. The best results were obtained in all subjects at input size fifty-four.

Table 1. Selecting optimal values for the SVM parameters

RBF_Sigma	Boxconstraint
0.9802	0.9658
1.5673	1.1059

Table 2 shows that the classification accuracy of the four imagery tasks reached 100% which is considered as a major progress considering the classification of EEG signals. The rest of the subjects also have 100% classification rate.

Table 2. The classification rate of imagery movement tasks for different subjects at input size = 54

Subject	Imagery Tasks			
	Left Hand	Right Hand	Foot	Tongue
Subject 1	100%	100%	100%	100%
Subject 2	100%	100%	100%	100%

To evaluate our system, we compare our results with those of [14]. We wanted to test our system using the same dataset. However, it was not applicable to get the data set used in [14] to use in our experiments as it is not available online. However we are testing for the same activities, the method mentioned in [14] uses also three different activities (Left hand, right hand and foot), so the dataset is expected to be similar. The comparison results are shown in Table 3, which reveals the outstanding performance of our system.

Table 3. Classification rate comparison

Motor Imagery Task / System	Left Hand	Right Hand	Foot	Tongue
Our System	100%	100%	100%	100%
Jiang Wang System	83.3%	82.1%	93.8%	

5 Conclusions

In this paper, EMD is applied to analyze EEG signals in four different motor imagery tasks. Seven subjects were used in our experiment. They imagined the movements of left hand, right hand, foot and tongue. An SVM classifier with RBF kernel was used to classify the recorded signals. The classification accuracy obtained, which reached 100% shows that the methodology used is very efficient for EEG signals classification. However, the classification performance on a huge number of samples is to be investigated in the future.

References

[1] Grabianowski (ed.): How brain-computer interfaces work (November 2007)
[2] Leeb, R., Friedman, D., Müller-Putz, G.R., Scherer, R., Slater, M., Pfurtscheller, G.: Self-paced (asynchronous) bci control of a wheelchair in virtual environments: a case study with a tetraplegics. In: Computational Intelligence and Neuroscience, p. 79642 (2007)
[3] Scherer, R., Lee, F., Schlogl, A., Leeb, R., Bischof, H., Pfurtscheller, G.: Eeg-based interaction with virtual worlds: A self-paced three class brain-computer interface. In: Computational Intelligence and Neuroscience (2007)
[4] Plant, K.A., Ponnapalli, P.V.S., Southall, D.M.: Mobile robots and eeg - a review. In: SGAI Conf., pp. 363–368 (2007)
[5] Satti, A.R., Coyle, D., Prasad, G.: Self-paced brain-controlled wheelchair methodology with shared and automated assistive control. In: IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (2011)
[6] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454, 903–995 (1971)
[7] Osuna, E., Freund, R., Girosi, F.: Support vector machines: Training and applications. Technical report, Cambridge, MA, USA (1997)
[8] Wang, J., Xu, G., Wang, L., Zhang, H.: Feature extraction of brain-computer interface based on improved multivariate adaptive autoregressive models. In: 2010 3rd International Conference on Biomedical Engineering and Informatics (BMEI), vol. 2, pp. 895–898 (October 2010)
[9] Herman, P., Prasad, G., McGinnity, T.M., Coyle, D.: Comparative analysis of spectral approaches to feature extraction for eeg-based motor imagery classification. IEEE Transactions on Neural Systems and Rehabilitation Engineering 16(4), 317–326 (2008)
[10] Jangraw, D.C., Sajda, P.: Feature selection for gaze, pupillary, and eeg signals evoked in a 3d environment. In: Proceedings of the 6th Workshop on Eye Gaze in Intelligent Human Machine Interaction: Gaze in Multimodal Interaction, GazeIn 2013, pp. 45–50. ACM, New York (2013)
[11] Hosni, S.M., Gadallah, M.E., Bahgat, S.F., AbdelWahab, M.S.: Classification of eeg signals using different feature extraction techniques for mental-task bci. In: International Conference on Computer Engineering Systems, ICCES 2007, pp. 220–226 (2007)

[12] Suleiman, A.-B.R., Fatehi, T.A.-H.: Features extraction techniques of eeg signal for bci applications, p. 5 (2011)
[13] Liang, H., Bressler, S.L., Desimone, R., Fries, P.: Empirical mode decomposition: a method for analyzing neural data. Neurocomputing 65-66, 801–807 (2005)
[14] Wang, J., Xu, G., Wang, J., Yang, S., Yan, W.: Application of hilbert-huang transform for the study of motor imagery tasks. In: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2008, pp. 3848–3851 (2008)
[15] Demir, B., Erturk, S., Gullu, M.K.: Hyperspectral image classification using denoising of intrinsic mode functions. IEEE Geoscience and Remote Sensing Letters 8(2), 220–224 (2011)
[16] Weng, B., Barner, K.: Optimal signal reconstruction using the empirical mode decomposition. EURASIP Journal on Advances in Signal Processing 2008(1), 845294 (2008)
[17] Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., Arnaldi, B.: A review of classification algorithms for eeg-based brain-computer interfaces. Journal of Neural Engineering 4(2), R1–R13 (2007)
[18] Lambert, M., Engroff, A., Dyer, M., Byer, B.: Empirical mode decomposition (December 2002)
[19] Rutkowski, T.M., Mandic, D.P., Cichocki, A., Przybyszewski, A.W.: Emd approach to multichannel eeg data - the amplitude and phase synchrony analysis technique. In: Huang, D.-S., Wunsch II, D.C., Levine, D.S., Jo, K.-H. (eds.) ICIC 2008. LNCS, vol. 5226, pp. 122–129. Springer, Heidelberg (2008)
[20] Brunner, C., Leeb, R., Müller-Putz, G., Schlögl, A., Pfurtscheller, G.: BCI Competition - Graz data set A (2008)