

Selecting the best features of EEG signals using CSP algorithm

Noran Elkafrawy, Doaa Hegazy and Sayed Fadel

Scientific Computing Department, Ain Shams University
Cairo, Egypt

Abstract

Brain Computer Interfaces (BCIs) is a powerful tool to assist people with disabilities. In this paper we focus on interpreting motor imagery tasks. We propose a feature extraction method based on Empirical Mode Decomposition (EMD), where The Electroencephalography (EEG) signal is decomposed into Intrinsic Mode Functions (IMFs) by the EMD algorithm and six statistical estimated parameters are calculated. Afterwards, Common Spatial Pattern (CSP) is applied to filter the feature vector and select the best features to overcome the curse of dimensionality problem. Then, the resultant features vector is fed to a Support Vector Machine (SVM) for classification. Promising results are obtained by testing the proposed model on the publicly available BCI competition 2008 dataset where a kappa result of 0.44 is achieved.

Keywords: Brain Computer Interface (BCI), Common Spatial Pattern (CSP), EEG Signals, Empirical Mode Decomposition (EMD), Support Vector Machine (SVM).

1. Introduction

Brain Computer Interfaces (BCI) have tremendous impact on the lives of people with disabilities. In accidents like stroke, head injury, spinal cord injury and multiple sclerosis, patients could suffer from complete body paralysis (NHS, 2018). This prevents them from using keyboards, mouse, or other means that persons without disabilities can use to interact with computers. BCIs can translate your thoughts into instructions that could be comprehended by computers. Whenever we think or move, our brains generate electrical signals that could be detected by scientists using special electrodes.

BCIs is very common to interface with humans, however several laboratories have managed to record some animals' signals, such as: monkey and rats. Monkeys were able to move cursors on the computer screen and were able to control a robotic arm so it could performs simple tasks. (Li et. al., 2017)

In this paper we are interested in the EEG modulated by motor imagery tasks (MI). We focus on extraction, selection and classification of best features calculated from EEG signals. Motor imagery is known to induce muscle activity and brain activity similar to that induced when performing the actual activity (Kim et. al., 2014). It is the mental representation of the performance without doing the action (Shenoy & Vinod, 2014).

Many applications use data from BCI systems to make life easier and more interesting, such as: Assistive technology, virtual reality, game controlling and robotics (Leeb et. al., 2007; Scherer et. al., 2007; Plant & Southall, 2007; Satti et. al., 2011). In these examples, authors are interested in understanding brain waves to be able to use them as inputs to other systems. Several studies have been made in this field with varying results.

We propose a method to classify EEG signals that consists of five steps: 1) Pre-processing of input signals to reduce the noise. 2) Extracting features using Empirical Mode Decomposition (EMD) (Huang et. al., 1998). 3) Calculating the estimated parameters for the output features (Lotte et. al., 2007). 4) Filtering the features using Common Spatial Pattern (CSP) **Error! Reference source not found.** (Mueller-Putz et. al., 1999). 5) Classifying the signals using Support Vector Machine (SVM) (Osuna et. al., 1997).

Experimental evaluation of the proposed method is applied to real human data and resulted in good results relative to the State-of-The-Art algorithms.

In section two, related work, where techniques used to solve the classification problem are presented. The proposed work with all its details is presented in section three. Data set used for experiments and the results obtained are discussed in section four. Finally, in section five the conclusions are drawn.

2. Material and methods

There are several techniques used for EEG feature extraction (Herman et. al., 2008; Hosni et. al., 2007; Jangraw et. al., 2013; Satti et. al., 2011; Wang et. al., 2010). Toka (Toka, 2011) gave an overview of the different methods to extract features from an EEG signal. These methods include Empirical Mode Decomposition (EMD) and Fourier Transform, which has Frequency Analysis (FFT), Space-Time-Frequency Analysis (STFT) and Time-Frequency Analysis (TFT).

Feature extraction methods are not restricted to Fourier Transform. Liang *et. al.* (Liang et. al., 2005) showed that Fourier-based methods are designed for stationary time series frequency analysis. Liang *et. al.* used Empirical Mode Decomposition (EMD) to analyze neural data coming from macaque monkeys. They identified high-frequency components as gamma band oscillations, and low-frequency components as average visual evoked potential (AVEP). They concluded that EMD might be a vital technique to analyze neural data.

EMD was first introduced by Huang *et. al.* (Huang et. al., 1998) as a signal decomposition method. It decomposes the signal into oscillatory modes called Intrinsic Mode Functions (IMFs). EMD does not require any pre-defined basis function to represent the signal as in Fourier or Wavelet methods.

Wang et. al (Wang et. al., 2008) used Hilbert-Huang transform method to analyze EEG signals and classify them into different classes. The used EEG signals consisted of different motor imagery tasks. The method used consisted of two steps: 1- EMD. 2- Hilbert spectral analysis. They also used BP Neural Network in classification (Amanpour & Erfanian, 2013). They classified three different tasks (left hand, right hand and foot) and reached a classification rate of 93.8%.

Ren et al. (Ren et. al., 2016) proposed an efficient feature extraction framework which combined hybrid feature extraction and selection methods. They applied different methods for feature extraction in an autoregressive model: Discrete Wavelet Transform, Wavelet Packet Transform and sample entropy. Afterwards, they select the relevant features to enhance the performance. Results were evaluated using class separability experiments on the EEG dataset of university of Bonn.

Bajaj and Pachori presented a method for classifying EEG signals in (Bajaj & Pachori, 2012). They used Empirical Mode Decomposition and considered the Intrinsic Mode Functions as a set of amplitude and frequency modulated signals (AM-FM). They

computed the two bandwidths from the analytic IMFs provided from the Hilbert transformation of IMFs. They used LS-SVM classifier and their best classification results for seizure and nonseizure EEG signals was 100%, using the second IMF with Morlet wavelet kernel function.

However, most researchers believe that feature selection is very important as it optimize time and performance by selecting only best feature for learning and classification. Alotaiby et al. (Alotaiby et. al., 2015) presented a survey on different channel selection algorithm. Their main purpose of channel selection process is to reduce computational complexity of any process performed on an EEG signal, reduce the overfitting that may arise by utilizing unnecessary channels and reduce the setup time of different applications. They mentioned the different methods for channel selection would include: time-domain analysis, power spectral estimation and wavelet transform. In addition to different approaches which include filtering, wrapping, embedded, hybrid and paper (human-based). The difference between these methods is described as follows:

1. Filtering:

Criterion used: Distance measure, information measure, dependency measure and consistency measure.

Pros: High speed, independent of classifier and scalability.

Cons: Low accuracy.

2. Wrapping:

Criterion used: Classification algorithm evaluate the channel subsets.

Pros: Good accuracy.

Cons: More complex than filtering and prone to overfitting.

3. Embedding

Criterion used: The criteria on selection is generated during learning process.

Pros: Interaction between channel selection and classification and less prone to overfitting.

4. Hybrid:

Criterion used: Combination of filtering and wrapping, independent measure selects the best subset and the mining algorithm select the final best subset.

Pros: Utilize both independent measure and a mining algorithm.

5. Human-based:

Criterion used: Well-trained observer.

Pros: Based on experience and used to refine the selection process.

Cons: Relatively low accuracy.

Error! Reference source not found. shows the 10-20 EEG electrodes placement used in this research from the left side and top of the head. These electrodes show different activities from different areas in the brain.

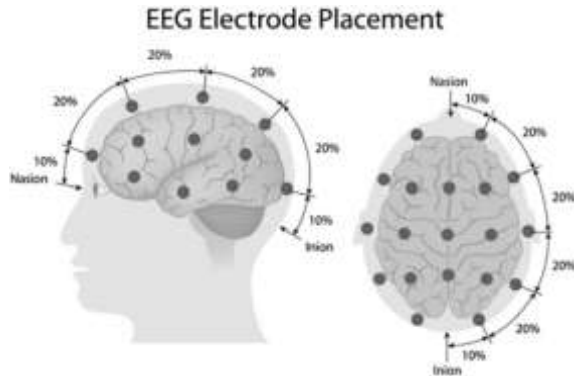


Fig. 1 International 10-20 system. The left image is a side view and the right image is a top view (B. graimann et. al., 2010)

Jangraw and Sajda (Jangraw et. al., 2013) used a virtual environment to explore which of the neural signals are evoked during the exploration of a dynamic, free-viewing 3D environment. They used a sequential forward floating selection (SFFS) and a hierarchical classifier that was adapted from the hierarchical discriminant component analysis (HDCA) to identify a small and robust set of features that was used to distinguish targets from distractors in this environment.

Tamil et. al. (Tamil et. al., 2008) reviewed the different algorithms used for feature extraction and classification. They concluded that the artificial neural network is the most popular for classification due to its accuracy. They compared it with back-propagation neural network, Bayesian neural network, polynomial neural network, fuzzy work neural network, chaos neural network, chaos theory, independent component analysis, Fourier transform, Laplacian eigenmaps method and Gabor transform.

D'albis et. al. (D'albis, 2012) developed a BCI system that predicts spelling upon motor imagery brain waves. They classified four classes: left hand, right hand, both hands and both feet. They applied a feature selection algorithm to select best features. They classified the features into two types: average features and evolution features. They classified the signals using linear discriminant analysis (LDA) and obtained a spelling rate of 2-3 character/minute.

Lotte et. al. (Lotte et. al., 2007) presented a review on classification algorithms for EEG based Brain Computer Interfaces. They divided the classification algorithms into five different categories: linear classifiers, neural networks, nonlinear Bayesian classifiers, nearest neighbor classifiers and

combination of classifiers (Figure 2). They concluded that SVM is very efficient regardless to the number of classes; this is because of its immunity to the curse of dimensionality. It is the most appropriate classifier to deal with feature vectors of high dimensionality.

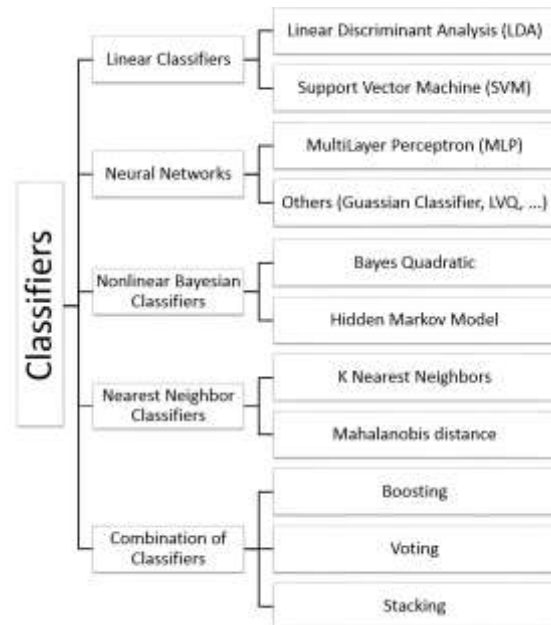


Fig. 2 Classification algorithms used to design BCI systems (El-Kafrawy et. al., 2014)

Medina-Slagdo et. al. (Medina-Salgado et. al., 2013) proposed a feature extraction method for EEG signals for imagining left hand and right hand. They used wavelet transform to decompose the signal in the spectral bands of interest. They conducted relevance analysis using fuzzy entropy algorithm to find the most important features in the training set. They used K-NN and SVM algorithms for classification and obtained an accuracy that reached 98.44%

3. Theory

Our system starts by reading a previously saved EEG signals dataset. This dataset contains signals with four different motor imagery tasks: imagining the movement for the left hand, right hand, both feet and tongue. Our main interest is to classify these signals into four different classes. In this paper feature extraction was the main focus. We used EMD to extract features (IMFs) and then calculated some statistical estimated parameters for each of these IMFs, Common Spatial Pattern (CSP) was applied to filter the resultant feature vectors. SVM classifier was then used with RBF kernel to classify the given

signals into four different classes. Figure 3 shows the diagram for the proposed model.

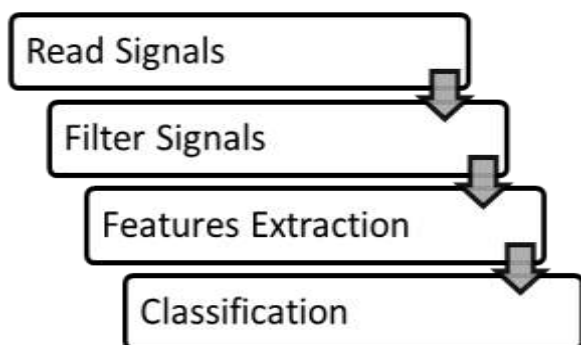


Fig. 3 Diagram of the proposed system

3.1 Pre-Processing

After sampling the signals with 250 Hz. A bandpass-filter between 0.5 Hz and 100 Hz was applied to remove the eye movement artifacts which was present in the dataset. The sensitivity of the amplifier was set to 100 μ V. An additional 50 Hz notch filter was enabled to suppress line noise.

3.2 Feature Extraction

The feature extraction process was very challenging to be able to select only the best features of each signal. The full process is shown in Error! Reference source not found. and the details will follow.

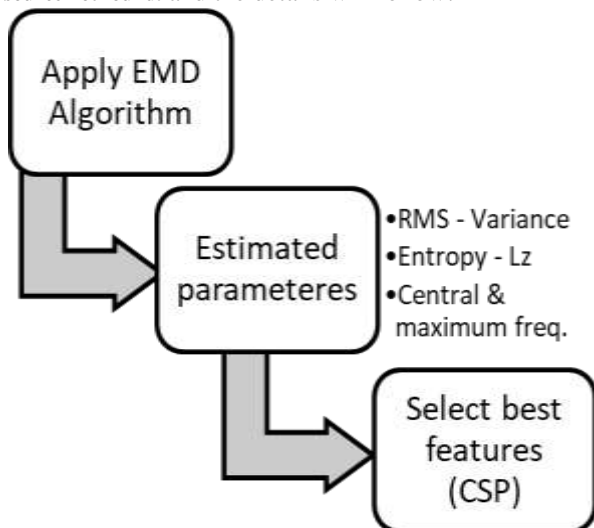


Fig. 4 Feature extraction

3.2.1 Apply EMD Algorithm (Empirical Mode Decomposition)

Empirical Mode Decomposition (Lambert & Byer, 2002) (EMD) is a method of breaking down a signal without leaving the time domain. It is useful in analyzing natural signals (like brain waves – EEG) (Weng & Barner, 2008). It

decomposes the signal into Intrinsic Mode Functions (IMFs) (Demir, 2011), which is easier to analyze.

An IMF is a function that has two properties: 1- it has only one extreme between zero crossings, and 2- it has a mean value of zero.

The EMD uses a process called sifting to decompose the signal into IMFs. This process is repeated until we get all the possible IMFs of the signal as shown in Figure 1

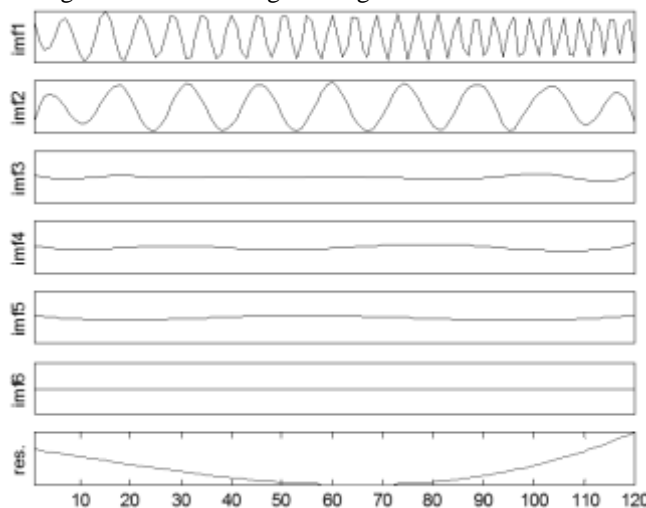


Figure 1. IMFs obtained using EMD

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 next step.

3.2.2 Calculate estimated parameters

For each of the first two modes obtained from the EMD and each EEG channel, six features were computed: (Diez, 2009)

- Root Mean Square (RMS),
- Variance,
- Shannon Entropy,
- Lempel-Ziv Complexity Value,
- Central Frequencies.
- Maximum Frequencies.

RMS and variance parameters are commonly used in BCI. Entropy was used as a different metric, as it measures the average amount of information from a measurement. LZ quantifies the complexity of a signal analyzing its spatio-temporal patterns. The central and maximum frequencies were used as descriptors of each IMF.

3.2.3 Select the best features (CSP)

The above method will result in very large feature vectors. Each vector will contain 96 parameters (2 IMFs x 6 parameters x 8 channels). As a result, we need to filter these vectors to get the best features and avoid the curse-of-dimensionality. Common Spatial Pattern is used to filter these feature vectors.

In Figure 2, we illustrate how CSP is used inside our system. CSP is considered the most popular algorithm in BCI field for learning spatial filters for oscillatory processes using the frequency band and time window *Error! Reference source not found.* Mueller-Putz et. al., 1999)

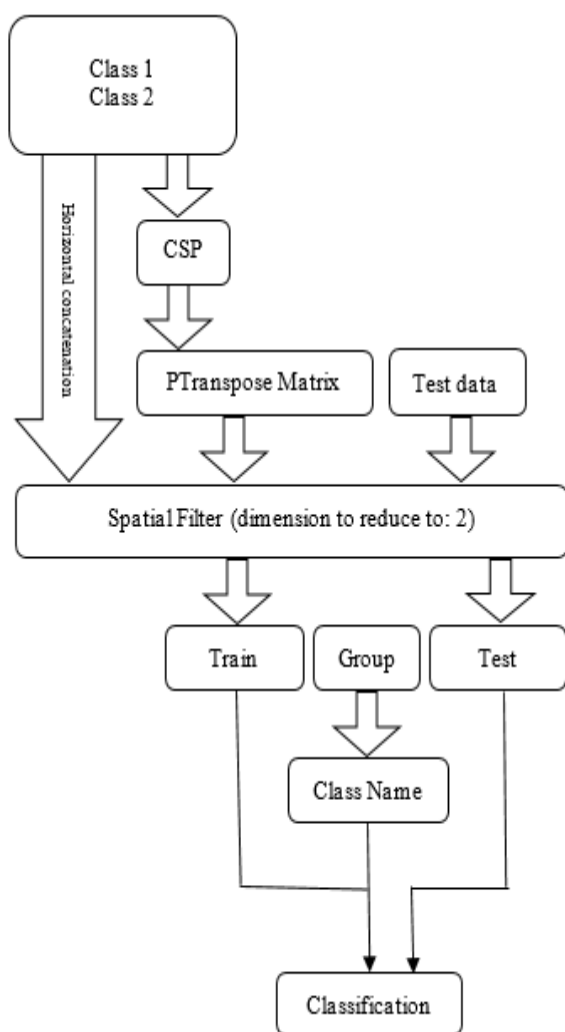


Figure 2 process showing how the CSP was used inside our system

3.3 Classification: Support Vector Machine

SVM is the most appropriate classifier to deal with high dimensional feature vectors (Grabianowski, 2007). There are several kernel functions that could

be used with SVM. The Guassian or Radial Basis Function (RBF) kernel are the most used in BCI classification tasks (Herman et. al., 2008). The corresponding SVM is known as Gaussian SVM or RBF SVM.

4. Results and Discussion

4.1 Data Set

In this work, we used the publicly online available BCI competition 2008 dataset (Brunner et. al., 2008). The data set is consisted of 9 subjects. Each of these subjects has four different motor imagery tasks: Imagining the movement of the left hand (class 1), right hand (class 2), both feet (class 3), and tongue (class 4). Each subject recorded two sessions on two different days.

Each of the recorded sessions consists of 6 runs, these runs are separated by short breaks. Each run consists of 48 trials, 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 shown in Figure 3.

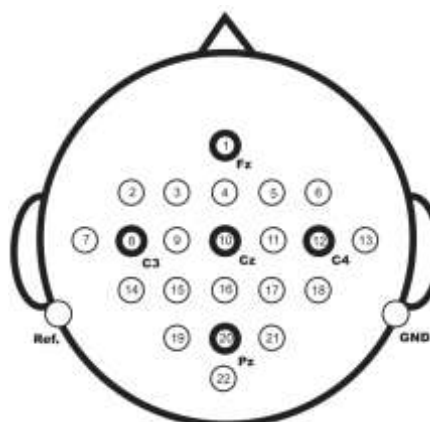


Figure 3. Electrode montage (Brunner et. al., 2008)

Figure 4 illustrates the paradigm of the process.

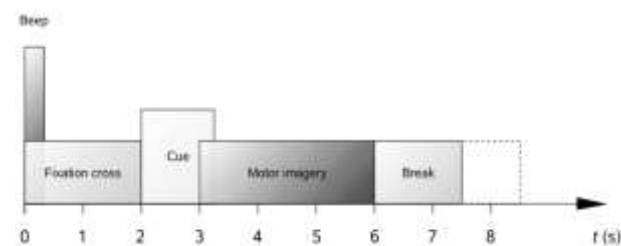


Figure 4. Timing scheme of the paradigm (Brunner et. al., 2008)

4.2 Performance Evaluation: Cohen's kappa

Several measures of performance have been proposed in BCI, such as; accuracy of classification, the most common is the accuracy of classification (the percentage of correctly classified feature vectors) and the Kappa coefficient. We consider the kappa coefficient.

Cohen's kappa measures the agreement between two raters.

A possible interpretation of Kappa

- Poor agreement = Less than 0.20
- Fair agreement = 0.20 to 0.40
- Moderate agreement = 0.40 to 0.60
- Good agreement = 0.60 to 0.80
- Very good agreement = 0.80 to 1.00

4.3 Experimental Evaluation

Different researchers had different results using the same dataset described in our paper and slightly different algorithms.

K. K. Ang -from the Institute of Infocomm Research Agency for Science, Technology and Research Singapore- (Ang et. al., 2012) reached **kappa = 0.57** by removing EOG with bandpass filter and extracting features with a filter bank CSP using mutual information rough set reduction (MIRSR) and Naïve Bayes Parsen window classifier.

L. Guangquan -from the School of Mechanical Engineering, Shanghai Jiao Tong University, China- reached **kappa = 0.52** using bandpass filters in different frequency bands then removing EOG, extracting features using Common Spatial Subspace Decomposition (CSSD).

D. Coyle -from the Intelligent Systems Research Centre, School of Computing and Intel Systems, UK- reached **kappa = 0.3** by applying CSP on spectrally filtered neural time series prediction preprocessing (NTSPP)

Our system reached **kappa = 0.44** which is relatively good value and above the chance level.

5. Conclusions

In this paper, EMD was applied to extract features, specific parameters were calculated from these features and CSP filter was used to reduce the dimension of features. An SVM classifier with RBF kernel was used to classify the output features. A moderate agreement value was obtained (0.44) which is above the chance level and better than four other methods using the same database.

More enhancements in the accuracy will be done as a future work.

References

[1] A. El-aal S, Ramadan R and Ghali N, "EEG Signals of Motor Imagery Classification Using Adaptive Neuro-Fuzzy Inference System", Proceedings of the 7th World Congress on Nature and Biologically Inspired Computing (NaBIC2015)., 11, 2016., pp. 105-116. Springer. (2016)

[2] Alotaiby T, El-Samie FEA, Alshebeili SA and Ahmad I, "A review of channel selection algorithms for EEG signal processing", EURASIP Journal on Advances in Signal Processing., Aug, 2015. Vol. 2015(1), pp. 66. (2015)

[3] Amanpour B and Erfanian A, "Classification of brain signals associated with imagination of hand grasping, opening and reaching by means of wavelet-based common spatial pattern and mutual information", In Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE., July, 2013. , pp. 2224-2227. (2013)

(Ang et. al., 2012) Ang KK, Chin ZY, Wang C, Guan C and Zhang H, "Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b.", Frontiers in Neuroscience. Vol. 2012;6:39 (2012)

[5] b. graimann, b. allison and g. pfurtscheller, "brain-computer interfaces" springer. (2010)

[6] Bajaj V and Pachori R, "Classification of Seizure and Nonseizure EEG Signals Using Empirical Mode Decomposition", Information Technology in Biomedicine, IEEE Transactions on. Vol. 16(6), pp. 1135-1142. (2012)

[7] Brunner C, Leeb R, Müller-Putz G, Schlögl A and GertPfurtscheller, "BCI Competition - Graz data set A", journal. (2008)

[8] D'albis T, Blatt R, Tedesco R, Sbattella L and Matteucci M, "A Predictive Speller Controlled by a Brain-computer Interface Based on Motor Imagery", ACM Trans. Comput-Hum. Interact. New York, NY,

USA, October 2012. Vol. 19(3), pp. 20:1-20:25. ACM. (2012)

[9] Demir B, Erturk S and Gullu M, "Hyperspectral Image Classification Using Denoising of Intrinsic Mode Functions", *Geoscience and Remote Sensing Letters*, IEEE. Vol. 8(2), pp. 220-224. (2011)

[10] Diez P, Mut V, Laciari E, Torres A and Avila E, "Application of the empirical mode decomposition to the extraction of features from EEG signals for mental task classification", In *Engineering in Medicine and Biology Society*, 2009. EMBC 2009. Annual International Conference of the IEEE., Sept, 2009., pp. 2579-2582. (2009)

[11] El-Kafrawy NM, Hegazy D and Tolba MF, "Features Extraction and Classification of EEG Signals Using Empirical Mode Decomposition and Support Vector Machine", In *Advanced Machine Learning Technologies and Applications - Second International Conference, AMLTA 2014*, Cairo, Egypt, November 28-30, 2014. Proceedings. , pp. 189-198. (2014)

[12] Grabianowski E, "How brain-computer interfaces work". (2007)

[13] Herman P, Prasad G, McGinnity T and Coyle D, "Comparative Analysis of Spectral Approaches to Feature Extraction for EEG-Based Motor Imagery Classification", *Neural Systems and Rehabilitation Engineering*, IEEE Transactions on., Aug, 2008. Vol. 16(4), pp. 317-326. (2008)

[14] Hosni S, Gadallah M, Bahgat S and AbdelWahab M, "Classification of EEG signals using different feature extraction techniques for mental-task BCI", In *Computer Engineering Systems*, 2007. ICCES '07. International Conference on., pp. 220-226. (2007)

[15] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N-C, Tung CC and Liu HH, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 454(1971), pp. 903-995. The Royal Society. (1998)

[16] Jangraw DC and 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*. New York, NY, USA , pp. 45-50. ACM. (2013)

[17] Kim J, Lee B-H, Suk Lee H, Ho Shin K, Ju Kim M and Son E, "Differences in Brain Waves of Normal Persons and Stroke Patients during Action Observation and Motor Imagery", *Journal of physical therapy science.*, 02, 2014. Vol. 26, pp. 215-8. (2014)

[18] Lambert Engroff D and Byer, "Empirical mode decomposition". (2002)

[19] Leeb R, Friedman D, Slater M and Pfurtscheller G, "BrainPlay'07: A tetraplegic patient controls a wheelchair in virtual reality", In *BRAINPLAY 07 Brain-Computer Interfaces and Games Workshop at ACE (Advances in Computer Entertainment) 2007*. (2007)

[20] Li L, Xu G, Zhang F, Xie J and Li M, "Relevant Feature Integration and Extraction for Single-Trial Motor Imagery Classification", *Frontiers in Neuroscience*. Vol. 11, pp. 371. (2017)

[21] Liang H, Bressler SL, Desimone R and Fries P, "Empirical mode decomposition: a method for analyzing neural data", *Neurocomputing* 65–66 (2005) 801–807. (2005)

[22] Lotte F, Congedo M, Lécuyer A, Lamarche F and Arnaldi B, "A review of classification algorithms for EEG-based brain-computer interfaces.", *Journal of Neural Engineering*. Vol. 4(2), pp. R1-R13. IOP Publishing. (2007)

[23] Medina-Salgado B, Duque-Munoz L and Fandino-Toro H, "Characterization of EEG signals using wavelet transform for motor imagination tasks in BCI systems", In *Image, Signal Processing, and Artificial Vision (STSIVA), 2013 XVIII Symposium of.*, Sept, 2013. , pp. 1-4. (2013)

[24] Mueller-Putz Guller-Gerking J, Pfurtscheller G and Flyvbjerg H, "Designing optimal spatial filters for single-

trial EEG classification in a movement task", *Clinical Neurophysiology*, 2017/11/30, 1999. Vol. 110(5), pp. 787-798. Elsevier. (1999)

[25] NHS, "Paralysis". (2018)

[26] Osuna E, Freund R and Girosi F, "Support Vector Machines: Training and Applications". Thesis at: Massachusetts Institute of Technology. Cambridge, MA, USA (1997)

[27] Plant KA, Ponnappalli PVS and Southall DM, "Mobile Robots and EEG - A Review", In SGAI Conf., pp. 363-368. (2007)

[28] Ren W, Han M, Wang J, Wang D and Li T, "Efficient feature extraction framework for EEG signals classification", In 2016 Seventh International Conference on Intelligent Control and Information Processing (ICICIP), Dec, 2016., pp. 167-172. (2016)

[29] Satti AR, Coyle D and 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)

[30] Scherer, R.; Lee, F. Y.; Schlögl, A.; Leeb, R.; Bischof, H.; Pfurtscheller, G EEG-based interaction with virtual worlds: A self-paced three class Brain-Computer Interface, In BRAINPLAY 07 Brain-Computer Interfaces and Games Workshop at ACE (Advances in Computer Entertainment) 2007. (2007)

[31] Shenoy HV and Vinod AP, "An iterative optimization technique for robust

channel selection in motor imagery-based Brain Computer Interface", In 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Oct, 2014., pp. 1858-1863. (2014)

[32] Tamil EM, Radzi HM, Idris MYI and Tamil AM, "A Review on Feature Extraction & Classification Techniques for Biosignal Processing (Part II: Electroencephalography)", 4th Kuala Lumpur International Conference on Biomedical Engineering 2008., In 4th Kuala Lumpur International Conference on Biomedical Engineering 2008., pp. 113-116. IFMBE Proceedings. (2008)

[33] Toka Abdul-Hameed Fatehi A-BRS, "Features Extraction Techniques of EEG signal for BCI Applications", journal., pp. 5. (2011)

[34] Wang J, Xu G, Jiang Wang, Yang S and Yan W, "Application of Hilbert-Huang transform for the study of motor imagery tasks", In Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE., pp. 3848-3851. (2008)

[35] Wang J, Xu G, Wang L and Zhang H, "Feature extraction of brain-computer interface based on improved multivariate adaptive autoregressive models", In Biomedical Engineering and Informatics BMEI, 2010 3rd International Conference on., Oct, 2010. Vol. 2, pp. 895-898. (2010)

[36] Weng B and Barner K, "Optimal Signal Reconstruction Using the Empirical Mode Decomposition", *EURASIP Journal on Advances in Signal Processing*. Vol. 2008(1), pp. 845294. (2008)