Motor Imagery based Brain Computer Interface using Transform Domain Features

Ahmed M. ELbaz[1], Ahmed T. Ahmed[1], Ayman M. Mohamed[1], Mohamed A. Oransa[1], Khaled S. Sayed[1] and Ayman M. Eldeib[1]

1

***Abstract*— Brain Computer Interface (BCI) is a channel of communication between the human brain and an external device through brain electrical activity. In this paper, we extracted different features to boost the classification accuracy as well as the mutual information of BCI systems. The extracted features include the magnitude of the discrete Fourier transform and the wavelet coefficients for the EEG signals in addition to distance series values and invariant moments calculated for the reconstructed phase space of the EEG measurements. Different preprocessing, feature selection, and classification schemes were utilized to evaluate the performance of the proposed system for dataset III from BCI competition II. The maximum accuracy achieved was 90.7% while the maximum mutual information was 0.76 bit obtained using the distance series features.**

***Index Terms* - Brain Computer Interface (BCI), Distance Series (DS), Motor Imagery, Nonlinear Dynamical Modelling, Phase Space Reconstruction.**

I. Introduction

Brain Computer Interface (BCI) is a device, which provides a non-muscular communication pathway between the brain and the external environment. The current trends in BCI focus on neuroprosthetics applications, which aim to restore damaged hearing, sight and movement. Artificial prosthesis can be used to replace the impaired functions of nervous system and brain related problems as well as sensory organs [1]. Moreover, BCI systems could be used for more advanced nonmedical applications such as gamming.

There are four types of EEG signals that can be used in BCI systems; Visual Evoked Potentials (VEP) [2], P300 Evoked Potentials [3], Slow Cortical Potentials (SCPs) and sensory rhythm (motor imagery) . In this study, we deal with motor imagery signals. Motor imagery based BCI systems translate subject's intention into a control signal that may be used by an external device. Researchers have developed many algorithms to classify motor imagery signals like in Lemm et al. [4] where they utilized the accompanying EEG μ-rhythm perturbation in order to distinguish between EEG signals while Baig et al. [5] used self-organizing maps (SOM) method to classify motor imagery EEG signals. Lee et al. [6] employed nonnegative tensor factorization (NTF) and used Viterbi algorithm to classify between mental tasks. Zhou et al. [7] introduced a feature extraction method based on bispectrum to classify right and left motor imagery signals. On the other hand, Darvishi and Al-Ani [8] used continuous wavelet transform to extract features from the signals and Adaptive Neuron-Fuzzy Interface System (ANFIS) for classification while Hazrati and Erfanin [9] introduced the usage of an adaptive probabilistic (APNN) for classification of motor imagery EEG signals.

In this work, we extracted different sets of transform domain features. They are the magnitude of the discrete Fourier transform (DFT), the discrete wavelet transform (DWT) coefficients, distance series (DS) values, and invariant moments. They are calculated for the reconstructed phase space (RPS) of the EEG signals, where each features set was utilized with different feature selection methods and different classifiers to obtain a system setting that enhances the performance measures for the motor imagery based BCI system.

II. Methodology

A. Preprocessing

We applied an 8-30 Hz Butterworth bandpass filter of order ten to eliminate the signal artifacts. The classification accuracy is expected to increase since the selected band of frequencies contains all μ and β frequency bands, which are associated with the activation of the motor cortex area [4].

B. Features Extraction

In this paper, we used different transform domain features such as DFT, DWT and the DS values, which was calculated for the RPS of the EEG signals as well as the moment invariant features.

1) Discrete Fourier Transform (DFT) Method

Transforming the brain signals into their frequency domain is considered as one of the methods used to recognize different mental tasks based on the EEG data. Here, we utilized the magnitude of the discrete Fourier transform over μ and β bands to differentiate between the imaginations of left- and right-hand movements.

2) Discrete Wavelet Transform (DWT) Method

Time–Frequency methods such as the wavelet transform are suitable to analyze the non-stationary signals such as EEG. The discrete wavelet transform (DWT) analyzes the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detailed information [10].

In our work, the EEG signals have been decomposed into three levels with Daubechies 4 (db4) family. The approximate

[1]Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University, Giza, Egypt (e-mail: eldeib@cu.edu.eg).

978-1-4577-0220-4/16/$31.00 ©2016 IEEE

coefficients Ca3 and the detailed coefficients at each level Cd3, Cd2 and Cd1 were used as features for the classification stage.

3) Nonlinear Dynamical Modeling of the EEG signals

Nonlinear dynamical modeling of the EEG signals provides an attractor or phase space (PS) which is equivalent to the true attractor of the underlying dynamical system [11].

In this work, we used the time-delay embedding methodology for reconstructing the PS where the optimal time-delay (τ) and optimal embedding dimensions (m) were selected based on the first minimum of the mutual information and the false nearest neighbour algorithm respectively [12].

The optimal time-delay and the optimal embedding dimension provided by [13] were calculated for each signal inside a sliding window over the whole training dataset and averaged to give a time-delay of 3 samples which was used to estimate the optimal embedding dimension m in the same way to give $m = 9$. PS could be reconstructed as shown in the following equation:

$$Y(m) = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_M \end{bmatrix} = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(m-1)\tau} \\ x_2 & x_{2+\tau} & \cdots & x_{2+(m-1)\tau} \\ \dots & \dots & \dots & \dots \\ x_M & x_{N-m\tau} & \cdots & x_N \end{bmatrix} \quad (1)$$

Where $Y(m)$ is the RPS, M is the number of the RPS points and equals $N - (m-1)\tau$, N is the length of original EEG time series, m is the optimal embedding dimension, and τ is the optimal time-delay.

Distance Series (DS)

The DS was proposed by Sayed et al. [14] to map the multidimensional phase space trajectory into one dimensional space by calculating the Euclidean distance between every point in the RPS and the origin of the space which can be calculated using the following equation:

$$D_i = \sqrt{\sum_{j=1}^{m} (Y_{ij} - Y_o)^2}\,, \quad i = 1, 2 \dots M \quad (2)$$

Where D_i is the Euclidean distance, Y_{ij} is a point in the phase space and Y_o is the origin of the RPS.

The DS domain makes it possible to visualize the behavior of the RPS and the evolution of the trajectory points which reveals more information about the underlying dynamical system.

Moment Invariant Features

The moment invariant features calculated for the RPS are utilized to numerically characterize the shape of m-dimensional PS trajectory with moments invariant to orthogonal transformations [15].

The steps of calculating the invariant moments of the EEG signals was described in [13] where the second order moments M were calculated by the following equation:

$$M_{p1\dots pm} = \sum_{i=1}^{M} \dots \sum_{i=1}^{M} s_{i1}^{p_1} \dots s_{im}^{p_m} \rho(s) ds_1 \dots ds_m \quad (3)$$

Where $\rho(S)$ is probability density function, S_i is a column in the RPS, and p is the order of the moments and is given by $p = \sum_{j=1}^{m} p_j$.

Central Moments were obtained using equation (4) from the second order moments, and then the O matrix was constructed from the central moments using equation (5). The major minors for the O matrix were calculated to represent the moment invariant features of the RPS giving that the number of the moment invariant features equals to the number of the embedding dimension m of the RPS.

$$\mu_{p1\dots pm} = \sum_{i=1}^{M} \dots \sum_{i=1}^{M} (s_{i1} - \bar{s}_{i1})^{p_1} \dots (s_{im} - \bar{s}_{im})^{p_m} \rho(s) ds_1 \dots ds_m \quad (4)$$

Where $\bar{s}_1 = \frac{M_{1\dots 0}}{M_{0\dots 0}}, \dots, \bar{s}_m = \frac{M_{0\dots 1}}{M_{0\dots 0}}$

$$O = \begin{bmatrix} \mu_{2\dots 0} & \cdots & \mu_{1\dots 1} \\ \vdots & \ddots & \vdots \\ \mu_{1\dots 1} & \cdots & \mu_{0\dots 2} \end{bmatrix} \quad (5)$$

C. Feature selection

Student's t-test was used for feature selection while the Principal Component Analysis (PCA) was used for dimensionality reduction and each of the two methods was utilized with every feature set to obtain the setting that produces the best performance measures.

D. Dataset Description

We used dataset III from BCI competition II [16] in which three bipolar EEG channels were measured over C3, Cz, and C4 electrodes. Bandpass filtering was of 0.5 and 30 Hz cut-off frequencies and sampling rate of 128 Hz. EEG signals were recorded from a 25-years old normal female subject which was asked to imagine left and right hand movements while setting on a relaxing chair with armrests. The experiment contains seven runs, each one is 40 trials with few seconds for rest after each run. Each trial is of 9 sec length with the first two seconds quiet. At time t= 2 sec an acoustic indicator is raised to declare the beginning of the trial and at t=3 sec, a left or a right arrow was displayed on the screen as a cue. At the same time, the subject is asked to imagine moving a bar in the direction of the cue.

E. Classification

To assess the performance of the purposed algorithm using different features, a set of different classifiers was used: k-nearest neighbor (KNN), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and support vector machine (SVM) with linear kernel.

We followed the classification procedure proposed by Fang et al. [17] where the features were extracted for each EEG signal in a sliding window of length equal 2 sec starts at t=3 and ends at t=9 sec as shown in Fig. 1. The classifier was trained and tested at each time point during the 6 seconds of the EEG signal providing classification results at each point in time for each algorithm.

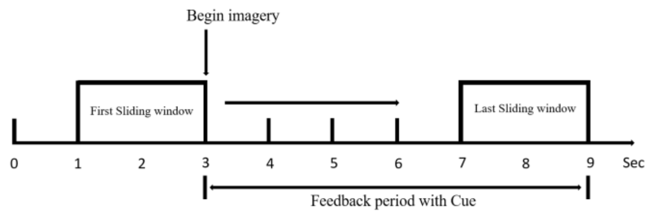

Figure 1. Sliding Window for dataset III from BCI competition II.

III. RESULT AND DICUSSION

To develop a certain algorithm that boosts the BCI performance measures for the utilized dataset, we performed 8 different experiments. In each experiment, one of the four feature sets was tested along with one of the two feature selection methods using the aforementioned classifiers.

Experiments were done using MATLAB on a PC with Intel® Core™ i7 2.60 GHz processor and 8 GB RAM. Experiments were evaluated in terms of the mutual information (MI) as well as the error rate (ERR) to compare our results with the other studies that used different algorithms for the same dataset. MI is the evaluation criterion for dataset III from BCI competition II and it is used to quantify the amount of information transfer between the subject's brain and the BCI system [18] while the ERR is the percentage of the misclassified signals.

Fig. 2 shows the time course of the MI obtained for the four experiments which showed the best performance as indicated in Table I. In addition, the time course of the ERR (ERR = 1 – Accuracy) is plotted in Fig. 3 for the same experiments.

Table I shows the maximum performance measures obtained using the extracted features, the corresponding feature selection method, and the classification technique. Moreover, the testing time for each algorithm is shown in Table I. It is clear that the experiment, which includes the DS values for the RPS, the PCA and the LDA classifier showed the best performance measures and obtained an accuracy of 90.71% and MI of 0.76 bit.

The time performance of experiment based on DFT was the best compared to the other experiments. However, the experiment based on DS obtained the maximum accuracy indicating that we need to develop faster algorithms by changing the classification procedure to reduce the overlap between successive windows (currently 255 /256 samples) to meet up with real time applications requirements.

Different algorithms were developed to increase the accuracy and boost the mutual information for dataset III from BCI competition II. To the best of our knowledge, Sayed [13] was the first publication that utilized the DS features for the RPS. In their work, they reconstructed a phase space for the EEG segments inside the sliding window with time-delay $\tau = 3$ and embedding dimension $m = 9$. Moreover, they used the student's t-test for feature selection as well as the KNN classifier with $k = 11$ for classification and they obtained maximum accuracy of 89.29% and maximum MI of 0.68 bit. In addition, the method developed by Fang et al. [17] used Amplitude Frequency Analysis of Phase Space (AFAPS) and Autoregressive Modeling of Phase Space (ARPS) while the method presented in Xu et al. [19] used wavelet-based features along with fuzzy support vector machine (FSVM). Zhou et al. [7] presented two experiments including bispectrum based features along with LDA, SVM and neural network (NN) classifiers. In addition, the methods purposed by Lemm et al. [4] used Morlet-Wavelets with Bayesian classifier.

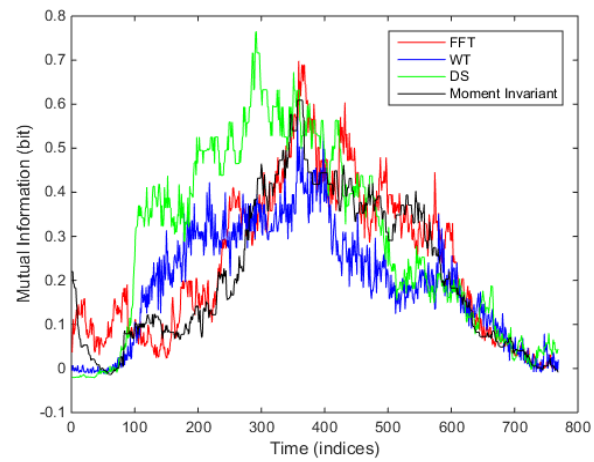

Figure 2. The MI of the best experiments.

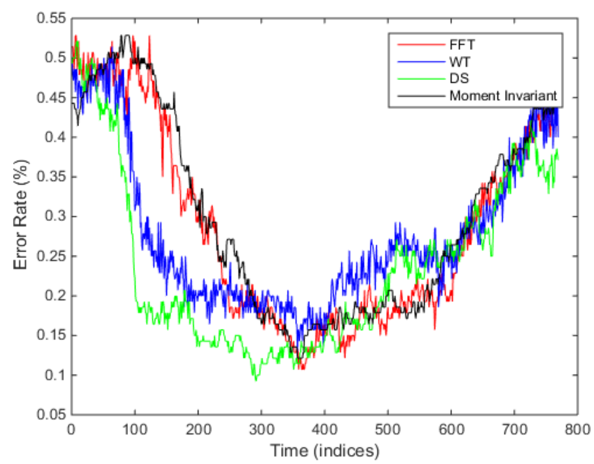

Figure 3. The ERR of the best experiments.

Table II shows a comparison between the proposed classification scheme and different algorithms developed for dataset III from BCI competition II. It is shown that the proposed algorithm obtained the best performance measures and outperformed the studies in comparison in terms of the MI, which was the evaluation criterion for the examined dataset.

IV. CONCLUSION

In this paper, we presented a BCI system based on transform domain features including the magnitude of the DFT, WT coefficients, DS of the RPS and invariant moments of the RPS. Student's t-test and PCA were used for feature selection and dimensionality reduction while different classifiers such as KNN, LDA, QDA and SVM with linear kernel were used. The best performance measures were obtained using the DS features along with the PCA and the LDA classifier. In addition, the maximum MI obtained here outperformed different state of the art algorithms that were developed to increase the MI for dataset III from BCI competition II.

TABLE I. THE BEST PERFORMANCE MEASURES OBTAINED.

Features	Max Acc (%)	Max MI (bit)	Time of max Acc (sec)	Feature selection	Classifier	Testing Time (sec)
Magnitude of FFT	89.28	0.672	5.80	Student`s t-test	KNN K=7	**0.01**
Coefficients of DWT	87.14	0.572	5.70	-	QDA	0.30
DS of the RPS	**90.71**	**0.764**	5.30	PCA	LDA	**0.03**
Invariant moments of the RPS	87.85	0.609	5.78	-	SVM with linear kernel	0.60

TABLE II. COMPARISON BETWEEN THE PURPOSED ALGORITHM AND OTHER STUDIES.

Method	Features	Max Acc (%)	Max MI (bit)	Time of max Acc (sec)	Classifier
Purposed algorithm	DS of the RPS	**90.71**	**0.76**	5.30	LDA
Sayed et al. [13]	DS of the RPS	89.29	0.68	5.84	KNN K=11
Fang et al. [17]	AFAPS+ARPS	90.71	0.67	5.76	LDA
Xu et al. [19]	Wavelet features	87.86	0.66	5.72	FSVM
Zhou et al. [7]	Higher-order statistics	90	0.64	-	NN
Zhou et al. [7]	Higher-order statistics	89.29	0.63	-	LDA
Lemm et al. [4]	Morlet-wavelets	89.29	0.61	7.59	Bayesian

REFERENCES

[1] H. S. Anupama, N. K. Cauvery, and G. M. Lingaraju, “Brain Computer Interface and Its Types - A Study,” *Int. J. Adv. Eng. Technol.*, vol. 3, no. 2, pp. 739–745, 2012.

[2] Y. Wang, X. Gao, B. Hong, C. Jia, and S. Gao, “Brain-Computer Interfaces Based on Visual Evoked Potentials,” *Ieee Eng. Med. Biol. Mag.*, vol. 27, no. 5, pp. 64–71, 2008.

[3] B. Rivet, A. Souloumiac, G. Gibert, and V. Attina, “‘P300 speller’ Brain-Computer Interface: Enhancement of P300 evoked potential by spatial filters BT - *16th European Signal Processing Conference, EUSIPCO 2008, August 25, 2008 - August 29, 2008*,” p. Swiss National Science Foundation SNSF; Interacti, 2008.

[4] S. Lemm, C. Schaefer, and G. Curio, “Probabilistic Modeling of Sensorimotor mu-Rhythms for Classification of Imaginary Hand Movements,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1077–1080, 2004.

[5] M. Z. Baig, Y. Ayaz, S. O. Gillani, M. Jamil, and M. Naveed, “Motor Imagery Based Eeg Signal Classification Using Self Organizing Maps,” *Sci. Int.*, vol. 27, no. 2, pp. 1165–1170, 2015.

[6] H. Lee, Y. Kim, A. Cichocki, and S. Choi, “Nonnegative tensor factorization for continuous EEG classification.,” *Int. J. Neural Syst.*, vol. 17, no. 4, pp. 305–317, 2007.

[7] S. M. Zhou, J. Q. Gan, and F. Sepulveda, “Classifying mental tasks based on features of higher-order statistics from EEG signals in brain-computer interface,” *Inf. Sci. (Ny).*, vol. 178, no. 6, pp. 1629–1640, 2008.

[8] S. Darvishi and A. Al-Ani, “Brain-Computer Interface Analysis using Continuous Wavelet Transform and Adaptive Neuro-Fuzzy Classifier,” *Eng. Med. Biol. Soc. 2007. EMBS 2007. 29th Annu. Int. Conf. IEEE*, pp. 3220–3223, 2007.

[9] M. K. Hazrati and A. Erfanian, “An online EEG-based brain-computer interface for controlling hand grasp using an adaptive probabilistic neural network,” *Med. Eng. Phys.*, vol. 32, no. 7, pp. 730–739, 2010.

[10] I. Dokare, “Classification of EEG Signal for Imagined Left and Right Hand Movement for Brain Computer Interface Applications,” *International Journal of Application or Innovation in Engineering & Management.* vol. 2014, pp. 291–294, 2014.

[11] N. H. Packard, J. P. Crutchfield, Farmer J. D., and S. R. S., “Geometry from a Time Series,” *Phys. Rev. Lett.*, vol. 45, no. 9, pp. 712–715, 1980.

[12] F. Takens, “Detecting strange attractors in turbulence,” Dyn. Syst. Turbul. Warwick 1980, vol. 898, pp. 366–381, 1981.

[13] K. S. Sayed, “Nonlinear Dynamical Modeling Features for Brain-Computer Interface” M. Sc. thesis, Biomedical Eng. Dept. . Cairo univ. Giza, Egypt, 2015.

[14] K. S. Sayed, A. F. Khalaf, Y. M. Kadah, "Arrhythmia classification based on novel distance series transform of phase space trajectories," *in Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* , vol., no., pp.5195-5198, 25-29 Aug. 2015.

[15] A. G. Mamistvalov, “n-Dimensional Moment Invariants and Conceptual Mathematical Theory of Recognition n-Dimensional Solids,” *IEEE trans. on pattern analysis and machine intelligence,* vol. 20, no. 8, pp. 819–831, 1998.

[16] “BCI Competition II.” [Online]. Available: http://www.bbci.de/competition/ii/.

[17] Y. Fang, M. Chen, and X. Zheng, “Extracting features from phase space of EEG signals in brain–computer interfaces,” *Neurocomputing,* vol. 151, pp. 1477–1485, 2015.

[18] A. Schlogl, C. Keinrath, and R. Scherer, “Information transfer of an EEG-based brain computer interface,” *1st InternationalIEEE EMBS Conference on Neural Engineering* vol.no. 1 0, pp. 641–644, 2003.

[19] Q. Xu, H. Zhou, Y. Wang, and J. Huang, “Medical Engineering & Physics Fuzzy support vector machine for classification of EEG signals using wavelet-based features,” *Medical Engineering & Physics 31 (2009)*,vol. 31, pp. 858–865, 2009.