Identification of Motor Imagery Movements from EEG Signals Using Automatically Selected Features in the Dual Tree Complex Wavelet Transform Domain

Syed Khairul Bashar*, Mohammed Imamul Hassan Bhuiyan

Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

Copyright ©2015 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract The decoding of human brain electrical functions by electroencephalogram (EEG) signal is the most important step in brain computer interface (BCI) based systems. So, in this paper, an automatic feature selection method has been proposed to classify imagery left and right hand movements from the EEG signals in the Dual Tree Complex Wavelet Transform domain. First, the EEG signals are decomposed into several bands of real and imaginary coefficients and then, some statistical features like Shannon entropy and variance have been calculated. These features are combined into a single feature space and after that optimal features have been selected automatically imposing some feature selection criteria from this combined feature space. The selected features have been shown to be promising to distinguish different kinds of EEG signals by statistical hypothesis testing (e.g., one way ANOVA) as well as graphical analysis (e.g., scatter plots, box plots). Finally, k-nearest neighbor based classifiers are developed using these selected features to identify left and right hand imagery movements. A mean accuracy of 90.00% is achieved in publicly available BCI competition II Graz motor imagery data set which is shown to be better than some existing techniques.

Keywords Electroencephalogram (EEG), Shannon Entropy, Dual Tree Complex Wavelet Transform, Feature Selection, kNN Classifier

1 Introduction

Brain computer interfacing (BCI) allows to control and operate computer aided systems by intent alone. The major objective of BCI is to assist disable people for their rehabilitation. BCI involves detection, analysis and classification of different types of motor imagery movements to implement real time control and communication. Electroencephalogram (EEG) signals are often used for BCI purpose since it can be implemented as a non-invasive system [1]. Analysis of EEG signals to decode human brain activity is essential for the implementation of BCI systems.

There are several categories of EEG-based BCI such as limb motor imagery classification [2], continuous arm movements direction detection [3], individual finger movement decoding [4], P300 evoked potential based character recognition [5] etc. One major category of BCI is the detection of motor imagery movements such as left and right hand movements. Various methods have been developed in the literature for classifying different types of arm movements. A wavelet-based common spatial pattern algorithm using low frequency features and Fisher linear discriminant classifier is developed to classify fast and slow hand movements in [6]. In [7], filter bank common spatial pattern is implemented with mutual information based feature selection and for the identification task, Naive Bayesian Parzen Window classifier is used. Wrist movement classification has been done by extracting gamma band features from wavelet packet transform and employing radial basis function classifier in [8]. Separability of EEG signals using adaptive auto regressive parameters is proposed in [9]. Time-frequency optimization is performed to classify left and right hand movements with reduced electrodes in [10]. For utilizing cross-channel dependency in BCI, multivariate empirical mode decomposition based classification method is presented in [11].

Since EEG data for research purpose can be acquired with varying experimental setup and conditions, BCI competition was held providing standard data sets to evaluate and compare different algorithms. The standard data sets were proved to be representative in motor imagery and were suitable for BCI research. Different approaches have been studied to classify motor imagery movements using the BCI competition II Graz motor imagery data set. Band passed EEG signals and power spectral density based linear discriminant analysis (LDA) is described in [12]. In [13], raw EEG signal incorporated with Hidden Markov Model (HMM) was presented by the same author. Adaptive Auto Regressive (AAR) model based features with Bayesian Graphical Network (BGN) and Multi Layer Perceptron are reported in [14]. Morlet wavelet is used to extract features from mu rhythms with Bayes quadratic classifiers in [15]. Wavelet coefficient based statistical features and fuzzy support vector machine (FSVM) classifier is described in [16]. Most recently, discriminative area selection method is implemented with fuzzy

Hopfield neural network (FHNN) classifier in [17].

The aim of this study is to assist developing BCI systems through identifying imagery hand movements by automatically extracting suitable features from EEG signals in the dual tree complex wavelet transform (DTCWT) domain. DTCWT has been widely used for image and video processing ([18], [19]), and recently, it has been used in the area of bio-medical signals ([20], [21]). But to the best of our knowledge, the DTCWT is being employed for the first time in the classification of imagery hand movements in conjunction with automatic feature selection. Since motor imagery movements occur in the low frequency EEG bands [22], to perform detail analysis on specific bands, DTCWT has been used as it is a rich level analysis tool than traditional discrete wavelet transform (DWT). The product of Shannon entropy and signal variance obtained from the various DTCWT bands are used as features. The ability of this value in discriminating imagery hand movements is demonstrated using various types of scatter plots, box plots and one way ANOVA analysis. An automatic selection method based on J_3 criteria is proposed for selecting the optimum feature vector. Left and right hand imagery movements are classified using these features employing a number of classifiers. The performance of these classifiers in detecting hand movements is obtained and compared with those of several existing techniques.

2 Description of the EEG database

The BCI competition II data set (motor imagery III) provided by Technical University of Graz has been used in this paper. This data was acquired from a normal subject while the subject sat in a chair with armrests and was trying to control a feedback bar by making imagery movements of left or right hands. Left and right cues were in random order [23].

Figure 1. Timing scheme of the experiment.

The experiment consists of 7 runs with 40 trials each. During each trial at $t=2s$ an acoustic stimulus indicates the beginning of the trial and a cross '+' was displayed for 1s. After this an arrow (left or right) was displayed at $t=3s$ as the cue. At the same time the subject was asked to move a bar into the direction of the cue which was controlled by adaptive auto-regressive parameters of channel C3 and C4. The EEG signal was filtered between 0.5 and 30 Hz while the sampling rate was 128 Hz. A detail description of the experimental set up can be found in [24]. Fig. 1 shows the timing scheme of the experimental setup while Fig. 2 presents the electrode/channel positions.

Figure 2. Electrode positions.

3 Dual Tree Complex Wavelet Transform

Dual tree complex wavelet transform is a recent enhancement to the discrete wavelet transform which has additional properties including nearly shift invariant and directionally selective in two and higher dimensions [25]. DTCWT is 2^d times redundant for any d dimensional signal as compared to DWT and offers directional information in six directions. Thus is it more efficient in time frequency localization of EEG signal.

Similar to positive/negative post-filtering of real subband signals, the idea behind dual tree approach is quite simple. DTCWT employs two real DWTs where the first DWT gives the real part of the transform while the second DWT gives the imaginary part. The analysis filter bank structure used to implement DTCWT is given in Fig. 3. Two real wavelet transforms use two different sets of filters which satisfy perfect reconstruction conditions. If square matrices $\mathbf{H}_h$ and $\mathbf{H}_g$ denote the two real DWTs, then the DTCWT can be represented as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_h \\ \mathbf{H}_g \end{bmatrix} \quad (1)$$

The inverse transform of $\mathbf{H}$ is given as,

$$\mathbf{H}^{-1} = \frac{1}{2} \begin{bmatrix} \mathbf{H}_h^{-1} & \mathbf{H}_g^{-1} \end{bmatrix} \quad (2)$$

If the vector $\mathbf{x}$ represents a real signal, then $\mathbf{w}_h = \mathbf{H}_h \mathbf{x}$ represents the real part and $\mathbf{w}_g = \mathbf{H}_g \mathbf{x}$ represents the imaginary part of the DTCWT.

Figure 3. Analysis filter bank of DTCWT.

When the dual-tree CWT is applied to a real signal, the output of the upper and lower filter banks gives the real and imaginary parts of the complex coefficients respectively.

Among several available wavelets, Farras wavelets [26] are used to perform DTCWT. Fig. 4 represents the EEG signals from C3 channel for left and right hand imagery movements and the corresponding second level DTCWT real and imaginary coefficients.

Figure 4. (a)-(b) EEG signal from C3 channels for left and right hand movement; (c)-(d) real coefficients of level 1 for left and right movements respectively; (e)-(f) imaginary coef. of level 1 for left and right; (g)-(h) real coef. of level 2 for left and right; (i)-(j) imaginary coef. of level 2 for left and right imagery movements.

4 Analysis in Dual Tree Complex Wavelet Transform Domain

The proposed method consists of first decomposing the EEG signal of each trial with dual tree complex wavelet transform into three levels and then extracting suitable features from different bands. After extracting suitable features, optimal features are selected automatically using J_3 criteria and finally kNN classifier is deployed for classifying hand movements.

4.1 Feature Extraction in DTCWT Domain

Since BCI competition II Graz data set was recorded with low pass filtering of 30 Hz, no preprocessing was needed to discard the unnecessary high frequency components. As discussed in [22], motor imagery activity occurs in low frequency EEG band signal. So to have in-depth view, wavelet transform is applied to the acquired EEG signal. Since forward transform of DTCWT gives two branches containing real and imaginary coefficients, it gives a rich way of analyzing EEG signals than DWT. The experiment was carried out by taking feedback from C3 and C4 channel while making Cz the reference. As a result, by applying DTCWT, EEG signal acquired from both C3 and C4 channel for a single trial is decomposed into three levels. If the original low passed signal is denoted by X which has 0.5 to 30 Hz frequency components, after first level decomposition, it provides Y_1 (16-30 Hz) and Z_1 (0.5-15 Hz). After second level decomposition, Z_1 leads to Y_2 (7.5-15 Hz) and Z_2 (0.5-7.5 Hz). So after three

levels of DTCWT, the four frequency parts are Y_1 (16-30 Hz), Y_2 (7.5-15 Hz), Y_3 (3.75-7.5 Hz) and Z_3 (0.5-3.75 Hz). Reconstructions of these components using the inverse DTCWT approximately correspond to the physiological EEG sub-bands delta, theta, alpha and beta respectively [27]. Since each frequency band gives both real and imaginary coefficients, we have total 4 real and 4 imaginary coefficient bands for each channel. From now on, the real and imaginary bands and associated levels will be denoted as RB_x and IB_x where 'x' is the level index.

The underlying dynamics of EEG signals is spread over various sub-bands in the frequency domain. To classify motor imagery movements, we need to extract more information in lower frequency bands of EEG signals and mu rhythms [15]. For this purpose, variance and Shannon entropy in different bands have been extracted as features.

4.1.1 Variance

Variance of a distribution is a measure of how widely values are dispersed from the average or mean value. It is the average squared distance between the mean and each item in the distribution. Variance is denoted by σ^2 where σ is defined as:

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}} \quad (3)$$

where N is the number of data points in a distribution, x_i is the i th sample of the distribution, $\bar{x}$ is the mean of N samples.

4.1.2 Shannon Entropy

Entropy is a measure of randomness which is a common concept in signal processing. Wavelet entropy can provide useful information about the underlying dynamical process associated with the signal [28]. The entropy E must be an additive cost function such that $E(0) = 0$ and $E(S) = \sum_i E(S_i)$, where S is the signal. In our study most common Shannon entropy has been used.

The Shannon entropy is defined as [29]

$$E(S) = - \sum_i S_i^2 \log(S_i^2) \quad (4)$$

where S_i are the coefficients of signal S in an orthonormal basis.

However, since variance and Shannon entropy both do not have good p-values individually in one way analysis of variance (ANOVA), they do not possess the ideal feature quality separately to discriminate left and right hand imagery movements. Also by using two features from each band, a total of $4 \times 2 \times 2 \times 2 = 32$ features are needed to be used for four levels from two channels which has very large search space for optimal feature selection. To overcome these problems, we have defined a combined feature space as the product of Shannon entropy and signal variance. This new feature space has two advantages. One is the separability of this combined feature space between two classes is better than the individuals (very small p-values of one way ANOVA analysis ensure this). The second advantage is that the length of feature vector is now reduced to half compared to separately using entropy and variance and thus keeping the search space smaller. So, the proposed method extracts features from Shannon entropy $\times$ variance plane for all four levels (i.e., eight bands) from two channels.

4.2 Automatic Feature Selection

Since the four level DTCWT of the EEG signal leads to eight coefficient bands for each channel, a total of 16 features are available from two channels. But how can we select the best feature combination? To combat this problem, we have used J_3 criteria [30] for selecting optimum feature combination. J_3 is a cost function for class separability measurement defined as

$$J_3 = \text{trace}\{\mathbf{S}_w^{-1}\mathbf{S}_m\} \quad (5)$$

Here $\mathbf{S}_w$ is the within class scatter matrix and $\mathbf{S}_m$ is the mixture scatter matrix defined as

$$\mathbf{S}_w = \sum_{i=1}^M P_i \mathbf{S}_i \quad (6)$$

$$\mathbf{S}_m = E[(\mathbf{x} - \boldsymbol{\mu}_0)(\mathbf{x} - \boldsymbol{\mu}_0)^T]$$

where, P_i is the priori probability of i^{th} class, M denotes total class number and $\mathbf{S}_i$ is the corresponding co-variance matrix given by

$$\mathbf{S}_i = E[(\mathbf{x} - \boldsymbol{\mu}_i)(\mathbf{x} - \boldsymbol{\mu}_i)^T] \quad (7)$$

where $\boldsymbol{\mu}_i$ is the mean of each class while $\boldsymbol{\mu}_0$ is the global mean. These two means are related by

$$\boldsymbol{\mu}_0 = \sum_{i=1}^M P_i \boldsymbol{\mu}_i \quad (8)$$

Sometimes between class scatter matrix $\mathbf{S}_b$ is used instead of $\mathbf{S}_m$ in (5). For M class problem, between class scatter matrix can be calculated as

$$\mathbf{S}_b = \sum_{i=1}^M P_i (\boldsymbol{\mu}_i - \boldsymbol{\mu}_0)(\boldsymbol{\mu}_i - \boldsymbol{\mu}_0)^T \quad (9)$$

J_3 criteria has the advantage that it is invariant under linear transform [31]. For a feature vector with a particular number of features, the feature elements yielding maximum value of J_3 are selected as the elements of the optimal feature vector. Here from available 16 features, 4 optimal features have been chosen automatically. The four features found using J_3 criteria are $RB_3(C3)$, $RB_4(C3)$, $RB_2(C4)$ and $IB_3(C4)$

Table 1. Results of ANOVA Analysis

Features	P-value from ANOVA
$RB_3(C3)$	2.4635e-16
$RB_4(C3)$	6.7988e-32
$RB_2(C4)$	3.9746e-04
$IB_3(C4)$	2.0261e-04

Table 1 shows the best possible features for a 4 element feature vector selected via the J_3 criterion and the corresponding p-values of one way ANOVA analysis. The hypothesis about the p-values is that the value, $p < 0.05$ indicates that at least one sample mean is significantly different than the other sample means statistically [32]. From Table 1, it is clear that these four features have very small p-values which indicates they can be used as good features.

Apart from one way ANOVA analysis, scatter plots and box plots are provided to further illustrate the classification quality of the features. Fig. 5 represents the scatter plots of

Figure 5. (a)-(b) Scatter plots of RB_3 and RB_4 during left and right hand imagery movements respectively for in C3 channel.

Figure 6. (a)-(b) Scatter plots of RB_2 and IB_3 during left and right hand imagery movements respectively for in C3 channel.

RB_3 and RB_4 for C3 channel, respectively whereas Fig. 6 presents the scatter plots of RB_2 and IB_3 for C4 channel, respectively. In Fig. 5 and Fig. 6, green hexagon and red square markers indicate feature values during left and right hand imagery movements respectively. The green and red markers have significantly different values indicating their variation during different imagery hand movements which states that they can be used as good features.

Fig. 7-10 present the box plots of $RB_3(C3)$, $RB_4(C3)$, $RB_2(C4)$ and $IB_3(C4)$ respectively for left and right hand imagery movements. The box plots indicate that the features have distinct values for the two specific movements and as a result they can be used in classifiers.

The scatter plots, box plots and the p-values of one way ANOVA indicates that $RB_3(C3)$, $RB_4(C3)$, $RB_2(C4)$ and $IB_3(C4)$ extracted using J_3 feature selection criteria from combined feature space of variance $\times$ Shannon entropy have distinguishable values for left and right hand motor imagery movements. In other words, the features have good between-class distance and small within-class variance in the feature vector space [31] and as a result they can be used as very good features to identify left and right hand motor imagery movements from EEG signals.

Figure 7. Boxplot of $RB_3(C3)$ for left and right hand imagery movements respectively.

Figure 8. Boxplot of $RB_4(C3)$ for left and right hand imagery movements respectively.

4.3 Classification Using kNN Classifier

k-Nearest Neighbors algorithm (kNN) is a non-parametric learning algorithm method used for classification. Among the various methods of supervised statistical pattern recognition, the Nearest Neighbor rule achieves consistently high performance, without a priori assumptions about the distributions from which the training examples are drawn [33].

In order to classify a sample trial vector X which has unknown class, kNN classifier ranks the sample trial's neighbors among the training trial vectors and uses the class labels of the k most similar neighbors to predict the class of the new test trial [34]. The classes of these neighbors are then weighted according to the similarity of each neighbor where the similarity index is the cosine value between two document vectors of Euclidean distance. The cosine similarity index is defined as

$$sim(X, D_j) = \frac{\sum_{t_i \in (X \cap D_j)} x_i \times d_{ij}}{\|X\|_2 \times \|D_j\|_2} \quad (10)$$

where X is the test or unknown trial; D_j is the j -th training trial; t_i is shared by both X and D_j ; d_{ij} is the weight for t_i in training sample D_j whereas x_i is the weight for t_i in X . The l_2 norm of X is defined as

$$\|X\|_2 = \sqrt{x_1^2 + x_2^2 + x_3^2 + \dots} \quad (11)$$

The number “ K ” decides how many neighbors influence

Figure 9. Boxplot of $RB_2(C4)$ for left and right hand imagery movements respectively.

Figure 10. Boxplot of $IB_3(C4)$ for left and right hand imagery movements respectively.

the classification. If $k = 1$, then the algorithm is simply called the nearest neighbor algorithm.

5 Experimental Results

The EEG data set has 140 trials each for left and right hand (total 280 trials) of 9 seconds length. Since the cue was given at $t=3$ sec, data segment after 3 seconds from C3 and C4 channels are used for classification. Only four features have been used to form the final feature vector. The train and test feature matrix has dimension of 140×4 which is fed to kNN classifier. The experiment was carried out using MATLAB 2013b [32] on Windows-7 32 bit platform having 1 GB RAM and 2.93 GHz Intel Core 2 Duo processor. Both training and testing data sets have been used in leave one out cross validation method. The accuracy has been calculated using the following equation:

$$Accuracy = \frac{\text{Correctly classified EEG epochs}}{\text{Total EEG epochs}} \times 100\% \quad (12)$$

First, the mean accuracy of the classification employing kNN classifier with different values of “ K ” is shown in Fig. 11. Here, the elements in the feature vectors are obtained by using the J_3 criterion. Note that, the best accuracies are obtained for $K=13, 9, 16$ and 16 for feature vectors with 4, 5,

6 and 7 elements, respectively (If the same accuracy occurs, only the lowest “K” value is selected). The distance parameter of the kNN classifier is selected as Euclidean distance [32].

Figure 11. Mean accuracy(%) vs different “K” values for various feature vector lengths.

Table 2 presents the feature elements and the corresponding best mean accuracy values of different feature combinations found by J_3 feature selection approach. If only two features are selected using J_3 criteria then the accuracy is very satisfying 87.86% which indicates that the combined feature space has significant distinguishable property for left and right hand imagery movements. When total number of selected features are 3, 4 and 5, then the mean accuracy is highest 90%. We have used four features for final comparisons (3 or 5 features can also be used similarly).

Table 2. Outcome of J_3 Criteria for Different Feature Combinations

Selected Feature Numbers	Selected Features	Mean Accuracy (%)
2	$RB_3(C3), IB_3(C4)$	87.86
3	$RB_3(C3), RB_4(C3), IB_3(C4)$	90.00
4	$RB_3(C3), RB_4(C3), RB_2(C4), IB_3(C4)$	90.00
5	$RB_3(C3), RB_4(C3), RB_2(C4), IB_3(C4), RB_4(C4)$	90.00
6	$RB_3(C3), RB_4(C3), RB_2(C4), RB_2(C4), IB_3(C4), RB_4(C4)$	89.64
7	$RB_1(C3), RB_3(C3), RB_4(C3), IB_4(C3), RB_1(C4), IB_3(C4), RB_4(C4)$	89.64

Table 3. Performance Evaluation of Different Classifiers

Classifier name	Classifier parameter	Mean Accuracy (%)
Discriminant Analysis	Mahalanobis distance	83.21
	Quadratic distance	86.43
	Linear distance	87.14
Naive Bayes	Gaussian Distribution	86.07
SVM	Radial Basis Kernel	87.50
kNN	Cosine distance	87.50
	Correlation distance	88.57
	Cityblock distance	88.57
	Euclidean distance	90.00

Table 3 provides the performance evaluation of different classifiers with varying parameters which indicates that kNN

classifier with Euclidean distance gives better accuracy than others.

Finally Table 4 compares the mean accuracy of the proposed method with several other methods. It can easily be concluded that using simple kNN classifier, our method provides good mean accuracy to classify left and right hand motor imagery movements than all the methods listed in table 4.

Table 4. Comparison of Mean Accuracy

Method	Proposed By	Classifier	Mean Accuracy (%)
PSD	Solhjoo et al. [12]	Mahalanobis distance	63.1
		Gaussian classifiers	65.4
		LDA	65.6
Raw EEG	Solhjoo et al. [13]	HMM	77.5
AAR	Tavakolian et al. [14]	Bayes Quadratic	82.86
		BGN	83.57
		MLP	84.29
Morlet wavelet	Lemm et al. [15]	Bayes Quadratic	89.29
Wavelet based features	Q. Xu et al. [16]	FSVM	87.86
Discriminative area selection	W. Hsu [17]	FHNN	83.10
Proposed method (DTCWT)		kNN	90.00

One may ask about the gain in performance achieved due to the use of J_3 criterion. Table 5 provides the mean accuracy values obtained by using 4 element feature vectors with various combinations of the elements. It also shows the mean accuracy obtained by using the feature vector whose elements are automatically selected by the J_3 criterion. From Table 5, it is clear that the mean accuracy obtained by using J_3 selection criterion is significantly higher than other combinations listed in the table and thus justifying the use of J_3 criterion.

Table 5. Mean Accuracy for Different Feature Combinations

Selected Features	Mean Accuracy (%)
$RB_1(C3), IB_1(C3), RB_2(C4), IB_2(C4)$	63.57
$RB_1(C4), IB_1(C4), RB_2(C3), IB_2(C3)$	77.14
$RB_2(C3), IB_2(C3), RB_3(C3), IB_3(C3)$	76.43
$RB_2(C4), IB_2(C4), RB_3(C4), IB_3(C4)$	71.43
$RB_2(C3), IB_2(C3), RB_3(C4), IB_3(C4)$	74.64
$RB_2(C4), IB_2(C4), RB_3(C3), IB_3(C3)$	78.21
$RB_1(C3), IB_1(C3), RB_1(C4), IB_1(C4)$	75.36
$RB_2(C3), IB_2(C3), RB_2(C4), IB_2(C4)$	83.93
$RB_3(C3), IB_3(C3), RB_3(C4), IB_3(C4)$	87.14
$RB_4(C3), IB_4(C3), RB_4(C4), IB_4(C4)$	76.07
4 features from J_3 criterion	90.00

Moreover, from Table 2 and Table 5, it can be said that, level 1 features derived from DTCWT (e.g. $RB_1(C3)$) are less suitable in classifying motor imagery movements than level 2 and 3. The underlying reason is that level 1 features correspond to 16-30 Hz frequency components whereas less than 15 Hz frequency elements (mu rhythms) are responsi-

ble for motor movements in neuro-physiological context. As a result, level 2 and 3 features perform relatively better in the hand movement classification problem. It can also be inferred from Table 1 that p-values of the selected features are much smaller for C3 channel than C4 and as a result, C3 channel may perform better than C4. Future works can be done to find more about the performance of individual channels to have further specific information about EEG dynamics. Also, feature selection criteria may be improved incorporating other between class inherent property of the selected features like orthogonality and mutual information.

6 Conclusions

In this paper a comprehensive method has been proposed to distinguish the left and right hand motor imagery movements which offers a promising support for an important application in BCI. EEG signals have been successfully classified by extracting features from a combined space formed by variance and Shannon entropy of EEG signals in the dual tree complex wavelet transform domain. Optimal combination of features have been selected automatically using J_3 criteria and justification of using these features has been provided with a number of scatter plots, box plots and one-way ANOVA analysis. Among various types of classifiers like SVM, kNN, Bayes classifier and LDA, kNN provides the highest accuracy of 90.00%. Finally the performance has been compared with several other recent methods available in EEG based BCI literature and shown to be superior to the others.

REFERENCES

- [1] Nicolas-Alonso, L.F.; Gomez-Gil, J. Brain computer interfaces, a review, *Sensors* 2012, 12, 1211-1279.
- [2] Yi, W.; Qiu, S.; Qi, H.; Zhang, L.; Wan, B.; Ming, D. EEG feature comparison and classification of simple and compound limb motor imagery, *J. Neuroeng. Rehabil* 2013, 10.
- [3] Woo, J.S.; Muller, K.R.; Lee, S.W. Classifying directions in continuous arm movement from EEG signals, *Brain-Computer Interface (BCI)*, 2015 3rd International Winter Conference on. IEEE, 2015, pp. 1-2.
- [4] Liao, K.; Xiao, R.; Gonzalez, J.; Ding, L. Decoding individual finger movements from one hand using human EEG signals, *PLoS ONE* 2014, 9.
- [5] Chen, Szi-Wen, and Yeh-Chi Lai. "A signal-processing-based technique for P300 evoked potential detection with the applications into automated character recognition." *EURASIP Journal on Advances in Signal Processing* 2014, no. 1 (2014): 1-10.
- [6] Robinson, N.; Vinod, A.P.; Ang, K.K.; Tee, K.P.; Guan, C.T. EEG-based classification of fast and slow hand movements using wavelet-CSP algorithm, *Biomedical Engineering, IEEE Transactions on* 2013, 60, 2123-2132.
- [7] Ang, K.K.; Chin, Z.Y.; Wang, C.; Guan, C.; Zhang, H. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b, *Frontiers in Neuroscience* 2012, 6.
- [8] Khan, Y.U.; Sepulveda, F. Brain-computer interface for single-trial EEG classification for wrist movement imagery using spatial filtering in the gamma band, *IET signal processing* 2010, 4, 510-517.
- [9] Neuper, C.; Schlogl, A.; Pfurtscheller, G. Enhancement of left-right sensorimotor EEG differences during feedback-regulated motor imagery, *Journal of Clinical Neurophysiology* 1999, 16, 373-382.
- [10] Yang, Yuan, Sylvain Chevallier, Joe Wiart, and Isabelle Bloch. "Time-frequency optimization for discrimination between imagination of right and left hand movements based on two bipolar electroencephalography channels." *EURASIP Journal on Advances in Signal Processing* 2014, no. 1 (2014): 1-18.
- [11] Park, Cheolsoo, David Looney, Alireza Ahrabian, and Danilo P. Mandic. "Classification of motor imagery BCI using multivariate empirical mode decomposition." *Neural Systems and Rehabilitation Engineering, IEEE Transactions on* 21, no. 1 (2013): 10-22.
- [12] Solhjoo, S.; Moradi, M. Mental task recognition: A comparison between some of classification methods, *BIOSIGNAL 2004 International EURASIP Conference*, 2004, pp. 24-26.
- [13] Solhjoo, S.; Nasrabadi, A.M.; Golpayegani, M.R.H. Classification of chaotic signals using HMM classifiers: EEG-based mental task classification, *Signal Processing Conference*, 2005 13th European. IEEE, 2005, pp. 1-4.
- [14] Tavakolian, K.; Vasefi, F.; Naziripour, K.; Rezaei, S. Mental task classification for brain computer interface applications, *Canadian Student Conference on Biomedical Computing*, 2006.
- [15] Lemm, S.; Schafer, C.; Curio, G. BCI competition 2003-data set III: probabilistic modeling of sensorimotor μ rhythms for classification of imaginary hand movements, *Biomedical Engineering, IEEE Transactions on* 2004, 51, 1077-1080.
- [16] Xu, Q.; Zhou, H.; Wang, Y.; Huang, J. Fuzzy support vector machine for classification of EEG signals using wavelet-based features, *Medical engineering & physics* 2009, 31, 858-865.
- [17] Hsu, W.Y. Brain-computer interface: The next frontier of telemedicine in human computer 251 interaction, *Telematics and Informatics* 2015, 32, 180-192.
- [18] Shi, F.; Selesnick, I.W. Video denoising using oriented complex wavelet transforms, *Acoustics, Speech, and Signal Processing*, 2004. Proceedings.(ICASSP'04). IEEE International Conference on. IEEE, 2004, Vol. 2, pp. ii-949.
- [19] Rahman, S.M.; Ahmad, M.O.; Swamy, M. Video denoising based on inter-frame statistical modeling of wavelet coefficients, *Circuits and Systems for Video Technology, IEEE Transactions on* 2007, 17, 187-198.
- [20] Das, A.B.; Bhuiyan, M.I.H.; Alam, S.S. Classification of EEG signals using normal inverse Gaussian parameters in the dual-tree complex wavelet transform domain for seizure detection, *Signal, Image and Video Processing* 2014, pp. 1-8.
- [21] Yang, H.; Guan, C.; Ang, K.K.; Wang, C.C.; Phua, K.S.; Yu, J. Dynamic initiation and dual-tree complex wavelet feature-based classification of motor imagery of swallow EEG signals. *Neural Networks (IJCNN)*, The 2012 International Joint Conference on. IEEE, 2012, pp. 1-6.
- [22] Lotte, F.; Congedo, M.; Lecuyer, A.; Lamarche, F. A review of classification algorithms for EEG-based brain-computer interfaces, *Journal of neural engineering* 2007, 4.

- [23] Pfurtscheller, G.; Neuper, C.; Schlogl, A.; Lugger, K. Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters, *Rehabilitation Engineering*, IEEE Transactions on 1998, 6, 316-325.
- [24] Graz description of data set III of BCI-competition 2003, Available online: [http:// www.bbc.de /competition/ii/](http://www.bbc.de/competition/ii/) (accessed on 15-06-2015).
- [25] Selesnick, I.W.; Baraniuk, R.G.; Kingsbury, N.G. The dual-tree complex wavelet transform, *Signal Processing Magazine*, IEEE 2005, 22, 123-151.
- [26] Wavelet Software. Available online: <http://eeweb.poly.edu/iselesni/WaveletSoftware/> (accessed on 21-06-2015).
- [27] Adeli, H.; Ghosh-Dastidar, S.; Dadmehr, N. A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy, *Biomedical Engineering*, IEEE Transactions on 2007, 54, 205-211.
- [28] Rosso, O.A.; Blanco, S.; Yordanova, J.; Kolev, V.; Figliola, A.; Schurmann, M.; Basar, E. Wavelet entropy: a new tool for analysis of short duration brain electrical signals, *Journal of neuroscience methods* 2001, 105, 65-75.
- [29] Wang, D.; Miao, D.; Xie, C. Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection, *Expert Systems with Applications* 2011, 38, 14314-14320.
- [30] Coelho, L.P.; Braga, D.; Dias, M.S.; Garcia-Mateo, C. An Automatic Voice Pleasantness Classification System Based on Prosodic and Acoustic Patterns of Voice Preference, *INTERSPEECH* 2011, pp. 2457-2460.
- [31] Theodoridis, S.; Koutroumbas, K. *Pattern Recognition*, fourth ed.; Academic Press Elsevier: Boston, 2009.
- [32] MATLAB. Available online: <http://www.mathworks.com/products/matlab/> (accessed on 25-06-2015).
- [33] Islam, M.J.; Wu, Q.J.; Ahmadi, M.; Sid-Ahmed, M.; others. Investigating the performance of naive-bayes classifiers and k-nearest neighbor classifiers, *Convergence Information Technology* 2007. International Conference on. IEEE, 2007, pp. 1541-1546.
- [34] Liao, Y.; Vemuri, V.R. Use of k-nearest neighbor classifier for intrusion detection, *Computers & Security* 2002, 21, 439-448.