Multi-class EEG classification of voluntary hand movement directions

Neethu Robinson[1]**, Cuntai Guan**[2]**, A P Vinod**[1]**, Kai Keng Ang**[2] **and Keng Peng Tee**[2]

[1] School of Computer Engineering, Nanyang Technological University, Singapore
[2] Brain Computer Interface Lab, Institute for Infocomm Research, Agency for Science, Research and Technology, Singapore

E-mail: neethu1@e.ntu.edu.sg

Abstract

Objective. Studies have shown that low frequency components of brain recordings provide information on voluntary hand movement directions. However, non-invasive techniques face more challenges compared to invasive techniques. *Approach*. This study presents a novel signal processing technique to extract features from non-invasive electroencephalography (EEG) recordings for classifying voluntary hand movement directions. The proposed technique comprises the regularized wavelet-common spatial pattern algorithm to extract the features, mutual information-based feature selection, and multi-class classification using the Fisher linear discriminant. EEG data from seven healthy human subjects were collected while they performed voluntary right hand center-out movement in four orthogonal directions. In this study, the movement direction dependent signal-to-noise ratio is used as a parameter to denote the effectiveness of each temporal frequency bin in the classification of movement directions. *Main results*. Significant ($p < 0.005$) movement direction dependent modulation in the EEG data was identified largely towards the end of movement at low frequencies ($\leqslant 6$ Hz) from the midline parietal and contralateral motor areas. Experimental results on single trial classification of the EEG data collected yielded an average accuracy of $(80.24 \pm 9.41)\%$ in discriminating the four different directions using the proposed technique on features extracted from low frequency components. *Significance*. The proposed feature extraction strategy provides very high multi-class classification accuracies, and the results are proven to be more statistically significant than existing methods. The results obtained suggest the possibility of multi-directional movement classification from single-trial EEG recordings using the proposed technique in low frequency components.

(Some figures may appear in colour only in the online journal)

1. Introduction

Brain–computer interfaces (BCIs) establish a communication system that bypasses the conventional neural pathways between the brain and muscles [1]. It translates the neural activity of the brain to control commands for driving external effectors. A major application of BCI is the rehabilitation of paralyzed patients with a cognitively intact brain but a non-functional spinal cord [1, 2]. BCI has the prospect to restore the movement capabilities of such patients, by interfacing neural activity with prosthetic devices. A major challenge in providing such movement control is the lack of finer information regarding movement related neural signals [3]. This information can be utilized in a BCI to provide a higher dimensional control command to drive an external device. The challenge is how well such information can be extracted from non-invasive electroencephalography (EEG) recordings [1, 3].

Neural oscillations over the sensorimotor area of the brain play an important role in voluntary movement execution and regulation of movement parameters [3, 4]. Many researchers

have used center-out movement studies to identify neural activity behind movement directions. Various studies have used invasive recording techniques, single-unit activity (SUA) [5], multi-unit activity (MUA) [6], localized field potential (LFP) [7], etc, to demonstrate direction dependent neural activities. Studies on primates using SUA and MUA have found cosine tuning of neural firing rates with movement direction, and have used this for direction decoding on a single trial basis [5, 6, 8]. A study [7] in monkeys using LFP, has reported that LFP in the frequency ranges $\leqslant$4 Hz, 6–13 Hz and 63–200 Hz provides movement direction information and the LFP spectra can thus be used for single trial movement direction decoding. Electrocorticography (ECoG), a partially invasive technique, was used in [9] to demonstrate the close correlation between cortical anatomy and multi-direction arm movement. The study reported significant spectral modulations in low (34–48 Hz) and high gamma (52–128 Hz) bands from the frontal and parietal lobe which were used for single trial decoding of hand movement direction. A non-invasive study using magnetoencephalography (MEG) in [10] reported power modulations in low pass filtered MEG activity ($\leqslant$3 Hz) and reported an average of 67% decoding accuracy across subjects. They also performed similar analysis with EEG and reported 55% decoding accuracy.

EEG recording electrical activity from the scalp during a neuronal activation has identified several movement related features [10–13]. The modulation of power in mu and beta bands causing movement event related synchronization and desynchronization [11] are well established features for single trial bilateral movement execution and imagery classification. Gamma band ($\geqslant$40 Hz) oscillations were studied by various researchers and showed enhanced spectral power before movement onset and hence were assumed to contribute to movement preparation [12]. Various studies [12, 13] reported that low gamma (30–50 Hz) and high gamma (50–150 Hz) bands show movement related power modulations prior to, during and toward the end of movement execution. However the study [13] reported movement direction dependent activation to be absent in the gamma activity, whereas the beta band (12.5–25 Hz) at the movement end provided movement direction information.

Most of these studies [3, 4, 7, 10, 14] reported that movement kinematics information is present in low frequency components of the neural signals. Notably [7] reported direction dependent tuning in LFPs <4 Hz and utilized the LFP amplitude spectra to decode hand movement direction in primates. The study in [10] that employed MEG and EEG made use of frequency components <3 Hz to extract information for movement parameter decoding and for single trial classifications. In this study, we present a novel signal processing technique to extract features from non-invasive EEG recordings for classifying voluntary hand movement directions. To the best of our knowledge, a similar investigation on low frequency EEG or applying similar signal processing techniques to effectively extract low frequency EEG direction related features has not been attempted in the literature to date. In our preliminary work [14], a regularized wavelet-common spatial algorithm (Reg. W-CSP) for the classification

of two movement directions was introduced. In this work, the proposed technique employs the Reg. W-CSP algorithm to extract low frequency components, a mutual information (MI)-based feature selection to select relevant features, and performs multi-class classification of four different directions. In this study, EEG data from healthy subjects were collected while they performed voluntary right hand center-out movement in four orthogonal directions, and the movement direction dependent signal-to-noise ratio (SNR) is used to investigate temporal, frequency and spatial components in the classification of movement directions. The effectiveness of the proposed technique is then investigated on the data collected and compared to existing methods.

2. Methods

2.1. Data recording

Brain electrical activity was recorded using a Neuroscan SynAmps 128 channel EEG amplifier during the experiment conducted at the Brain Computer Interface Lab at Institute for Infocomm Research. The data is sampled at 250 Hz, filtered at a lower cut-off frequency of 0.05 Hz and the signal bandwidth (highest frequency) is limited at 125 Hz by the acquisition system. EEG is recorded from seven healthy human male subjects. Electrooculography (EOG) is also recorded to remove ocular artifacts from the recorded EEG data.

2.1.1. Experimental task. The tasks involved two dimensional center-out horizontal movements in four orthogonal directions, indicated by figure 1(a), using the right hand while holding the MIT MANUS robot [15]. The robot recorded the position, velocity and force applied by the hand in two dimensional space for every sample time. A display screen placed in front of the subject provided the preparation, rest and movement cues. The experiment setup is shown in figure 1(d). The subjects were instructed to minimize eye movements to reduce EOG artifacts.

The experiment details, timeline and cue display screens are shown in figure 1(b). During recording, the home/rest period was indicated by an encircled cross at the center of the display. Each trial started with a rest period of about 4 s. The direction cue was presented by an empty circle in one of the four target positions, as shown in figure 1(b) for 'NORTH'. The task involved preparation for 2 s followed by movement, cued by the disappearance of the center circle. The maximum distance to be covered by the movement was 15cm, and the subjects were asked to complete the task within 0.5 s. The trial end was denoted by the appearance of a cross at the target. The trials in which subject failed to perform the correct movement, or complete the task within the required time, were flagged and the subject was notified of the error, denoted by feedback, shown in figure 1(b). The range of parameter measurements obtained from the robot is as follows: position (0 $\pm$ 0.15 m), velocity (0 $\pm$ 0.4 m s^{-1}) and force (0 to 10 N). The movement trajectory followed by each of the subjects is shown in figure 2. The dotted lines indicate the movement path for single trials and the trial-averages are indicated by the red curve.
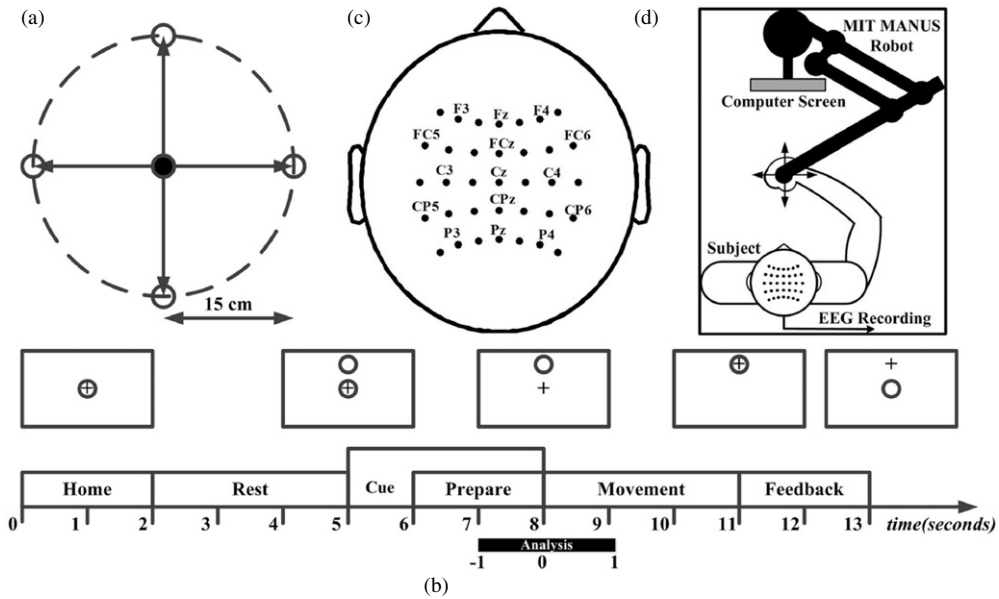
Figure 1. Experiment timeline and protocol. (a) The four orthogonal directions used for center-out movements in the experiment are shown. The dark circle at the center and four blank circles indicate the start point and target points, respectively. (b) Single trial experiment timeline for cue NORTH. The upper panel indicates the various cues presented on the display screen during the experiment. The bottom panel explains the cues along with the time periods. The time segment used for analysis is indicated by the 'analysis' bar. (c) The EEG sensor locations over sensorimotor cortex utilized in this study. (d) Experimental setup.
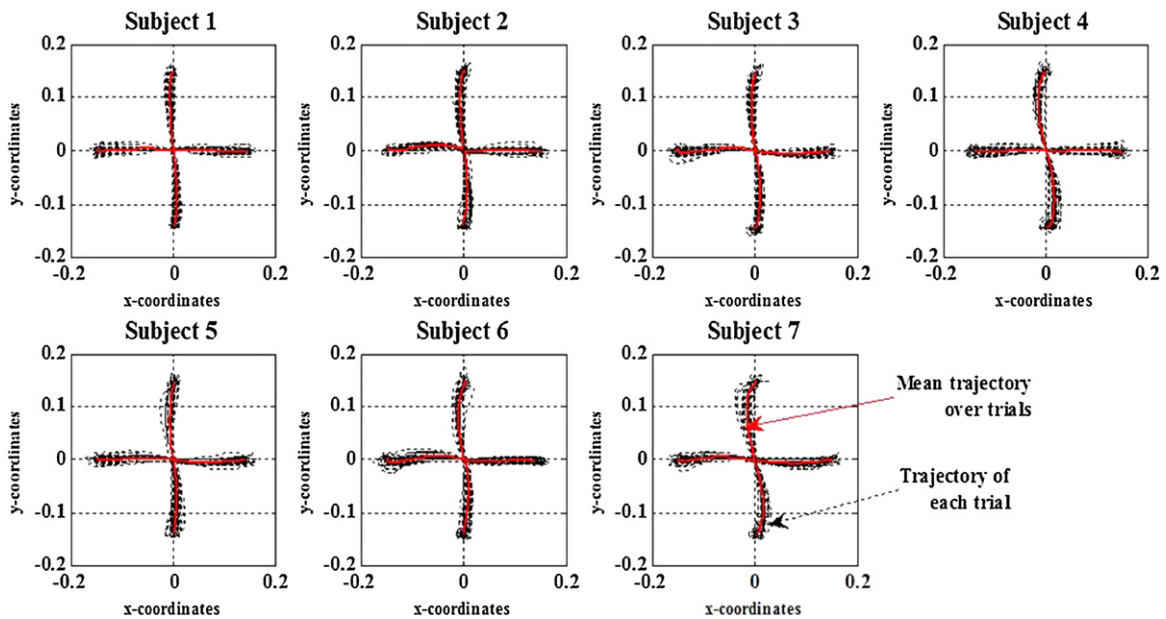
Figure 2. The hand movement trajectory for the subjects recorded by the MIT MANUS robot. The *x* and *y* coordinates are indicated in meters. It can be noted that in all cases the subjects follow the required trajectory for hand movements and for a maximum distance of 0.15 m.

2.1.2. Data pre-processing. The EEG data recorded from 35 sensors spanning the sensorimotor cortex, as shown in figure 1(c) are used for further analysis. The recorded data are low pass filtered at 96 Hz and notch filtered at 50 Hz to remove the line frequency. The ocular artifacts are removed from this data (f_{low}: 0.05 Hz to f_{high}: 96 Hz), using independent component analysis (ICA) [16]. The ICA components showing maximum correlation with actual recorded EOG are nullified. To remove muscular artifacts, we use a Laplacian spatial filter that accentuates localized activity whereby the diffused activity of EMG is suppressed.

The finite difference method reported in [25] is used to the derive Laplacian filter. The number of trials obtained is 160 for the first six subjects and 140 for the last subject, with an equal number of trial/movement direction for all subjects. As indicated in figure 1(b), the time segment extracted for analysis is 1 s before (−1 to 0 s) and after (0 to 1 s) the movement cue, to include the movement preparation, execution and post-movement periods. The extracted data are further used to extract informative direction-dependent features. Figure 3 shows the signal processing steps involved in the analysis of the data. The entire dataset is divided into training and testing
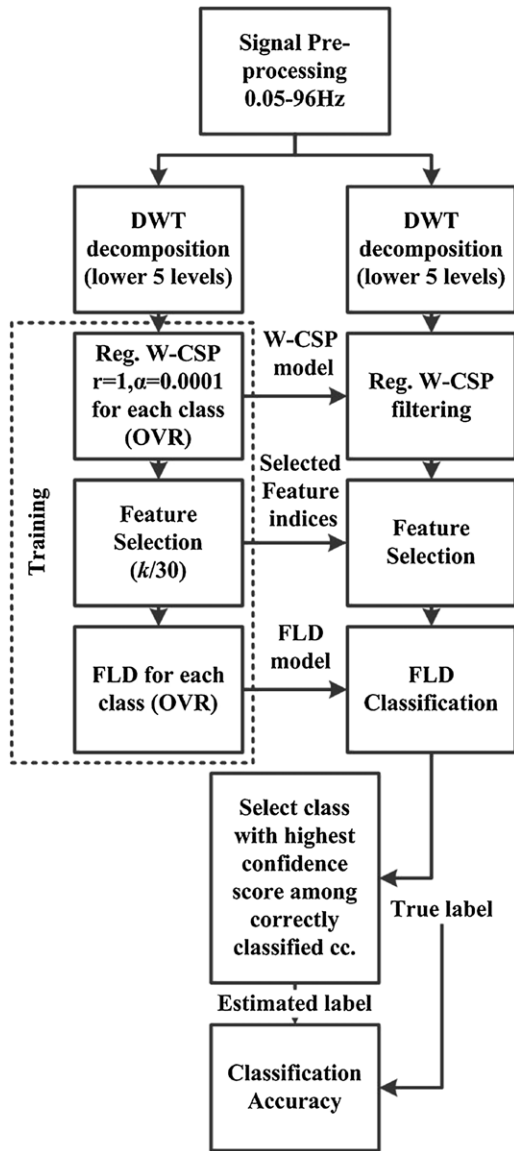
Figure 3. The block diagram for the signal processing algorithm used for the multi-class classification of hand movement directions in this study. The splitting of the data set into training and test folds during each cross-validation fold is indicated. The training data creates the Reg. W-CSP pattern, selects *k* features and models the FLD classifier. These are applied to the test set to calculate system performance.

sets in each of the cross validation folds. The calculation of spatial filters, feature selection and classifier modeling is done using the training dataset and the results are applied to the test data. Sections 2.2 and 2.3 explain the algorithm in detail.

2.2. Feature extraction using Reg. W-CSP

2.2.1. Wavelet decomposition of signal.
The time-frequency analysis of non-stationary EEG signals using Wavelet transform is a widely used feature analysis technique in BCI systems [17–19]. A common practice is to create a time-frequency distribution (TFD) of the recorded EEG using wavelet transforms and to define features using TFDs for various tasks. This approach has been applied in movement

and motor imagery studies [17, 18]. Another technique is to use wavelet decomposition for de-noising data and improving its SNR. In [19], W-CSP that used wavelet packets to decompose EEG data, followed by CSP filtering to analyze asynchronous BCI signals, was reported. In our analysis of multi-class data, we require decomposition of the signal into various subbands so as to localize the signal in the spectral domain and to reduce its non-stationarity before applying CSP for spatial localization. As literature reports the presence of direction dependent activity in low frequency regions, we require a feature extraction technique to focus on decomposing low frequency signals with high resolution. Hence, we use discrete wavelet transforms (DWT) for multi-resolution analysis of the pre-processed EEG data.

The orthogonal filter banks that span non-overlapping frequency subbands are constructed by DWT using orthonormal wavelet bases [20]. At each level of wavelet decomposition, the signal is half band low pass filtered and half band high pass filtered followed by sub sampling by a factor of 2, to obtain approximate and detailed wavelet coefficients. The approximate coefficient is further decomposed using orthogonal filters at the next level. This continues until maximum levels of decomposition given by $L_w = \log_2(T)$ are achieved. Thus we obtain L_w detailed coefficients and 1 approximate coefficient at the last level. The coefficients obtained at each level are separately reconstructed to obtain filtered subband signals. The frequency spanned by $L_w + 1$ subbands is given by,

$$\begin{aligned}
\left(2^{-i}f_{\text{high}}, 2^{-i+1}f_{\text{high}}\right), & \quad i = 1 \text{ to } L_w \\
\left(f_{\text{low}}, 2^{-i+1}f_{\text{high}}\right), & \quad i = L_w + 1.
\end{aligned} \tag{1}$$

The details of the steps following wavelet decomposition are explained in the following sections. As we have four classes of data, to perform multi-class data analysis and classification we use a one-versus-rest (OVR) approach in the following three stages, namely: feature extraction using Reg. W-CSP, feature selection using MI and classification using the Fisher linear discriminant (FLD) classifier. All these techniques use the binary class of data to generate their own functional models. The OVR approach creates four binary class problems, grouping one class against the remaining three classes. In the final stage, the classifier generates scores corresponding to the four OVR-FLD models and the class corresponding to the highest score is chosen as the estimated class.

2.2.2. Spatial filtering of subband signals.
The next step in our algorithm is to spatially filter the subband signals using the Reg. CSP approach. The subband signals in each trial are denoted by $Z^l \in \mathbb{R}^{C \times T}$, where $l = 1$ *to* L, where L is the number of subbands used for further analysis, T is the number of samples in each trial and C is the number of channels. For the binary class of data, class $\in a, b$, the CSP technique designs spatial filters that construct time series whose variances contain maximum discriminative information between classes [21].

By supervised decomposition of the signal, CSP creates a spatial filter, $W \in \mathbb{R}^{C \times C}$ which maximizes the variance for one particular class while minimizing it for the other class. An

optimization criterion is defined to determine the CSP filter [22]. For a particular subband l, spatial filter W^l is the matrix that extremizes the following objective function:

$$J(W^l) = \frac{W^{l^T} C_a^l W^l}{W^{l^T} C_b^l W^l} \qquad (2)$$

where $C_a^l = \frac{1}{n_a} \sum Z_a^l Z_a^{l^T}$ and $C_b^l = \frac{1}{n_b} \sum Z_b^l Z_b^{l^T}$ are the estimated covariance matrices of class a and b. Here n_a, n_b denote the number of trials in classes a, b and Z_a^l and Z_b^l denote trials belonging to their respective classes.

In the literature, several variants have been proposed to CSP, as it is prone to over-fit to the noise in the training data. An effective solution for this is regularizing CSP by introducing a penalty term in (2), so as to obtain a sparse spatial filter. In our work, we use spatial regularization that uses the spatial location of electrodes to define the penalty term and create spatially smooth filters [22]. In this work, by incorporating spatial regularization, the objective function in (2) takes the form,

$$J_a(W^l) = \frac{W^{l^T} C_a^l W^l}{W^{l^T} C_b^l W^l + \alpha.P(W^l)}. \qquad (3)$$

In (3), α is a positive constant that defines the level of spatial smoothness and $P(W^l)$ is the penalty function that measures the spatial smoothness of the filter. The function is defined as,

$$P(W^l) = W^{l^T}(D - G)W^l. \qquad (4)$$

In (4), G is a Gaussian kernel defined as, $= \exp(-\frac{\|v_i - v_j\|^2}{2r^2})$, where v_i *and* v_j are the spatial coordinates of electrodes i and j and r is the hyperparameter that defines the maximum distance between two electrodes so that they can be considered close. The matrix D is a diagonal matrix, whose diagonal element in each row corresponds to the sum of the respective row elements in G. Similar to (3), the objective function, $J_b(W^l)$ is also defined. Equation (3) can be defined as a constrained optimization problem in which the objective is to maximize $W^{l^T} C_a^l W^l$, while keeping $W^{l^T} C_b^l W^l + \alpha.P(W^l) = 1$. This optimization problem is solved using the Lagrange multiplier method as,

$$V_a^l = \left(C_b^l + \alpha(D - G)\right)^{-1}.C_a^l \qquad \text{and}$$
$$V_b^l = \left(C_a^l + \alpha(D - G)\right)^{-1}.C_b^l. \qquad (5)$$

The objective functions are maximized if the eigenvectors corresponding to the largest m ($= 3$ in this study) eigenvalues of V_a^l and V_b^l are chosen as the solution, $W^l \in \mathbb{R}^{2m \times C}$. We obtain the Reg. W-CSP filtered signal for subband l by scaling single trial data with the spatial filter, W^l. The single trial feature ($F^l \in \mathbb{R}^{2m}$) is obtained from the logarithm of normalized variances of Reg. W-CSP filtered data. In each of the four OVR class problems, the process is repeated for all the subbands to obtain a single trial feature set, $F \in \mathbb{R}^{2mL}$.

2.2.3. Selection of the most discriminative features. The MI-based best individual feature (BIF) selection technique has been successfully used, along with the filter bank CSP approach, in BCI studies to select the features that maximize the MI between classes and feature sets [23]. The selected feature set is proved to minimize the classification errors. We have $F = \{f_1, f_2, f_3, \ldots, f_{2\,mL}\}$ from which k ($<2\,mL$) best features are required to be selected for each of the four OVR classes. The MI between F and class ($\omega \in a, b$) is calculated as,

$$I(F; \omega) = H(\omega) - H(\omega|F). \qquad (6)$$

$H(\omega)$ and $H(\omega|F)$ denote the class entropy and conditional entropy functions respectively. Initially, the selected features are set as $\varnothing$. MI is calculated for the entire F and the feature with maximum MI is selected at each step. F is updated by eliminating the previously selected feature. The process is repeated until k features are selected.

2.3. Multi-class classification of Reg. W-CSP features

Similar to feature extraction, in this step we also use the OVR approach to perform multi-class classification. The FLD method [24] is used to classify the data. This technique determines a matrix, F, that maximizes the between class scatter, S_B, to within class scatter, S_W, of the features provided. It optimizes the objective function, as given by the mathematical expression in (7),

$$J_{\text{FLD}}(D) = \frac{D^T S_B D}{D^T S_W D}. \qquad (7)$$

The average cross validation accuracies obtained are reported as the performance measures.

2.4. Movement direction dependent SNR

In this section we attempt to determine the extent to which the wavelet domain analysis of EEG signals contributes to movement direction information. Here the parameter SNR of movement direction is used to identify the informative time-frequency regions of the signal. This type of analysis has been widely used in LFP studies [7, 8] to measure the direction tuning strength of neuron populations. In this study, the SNR is calculated for every sample point in all the levels of wavelet decomposition, using the reconstructed subband signals in each level. We calculated the direction tuning curves as the trial-averages of the signal corresponding to each direction:

$$\text{SNR} = \frac{\sigma_s^2 - \sigma_b^2}{\sigma_n^2}. \qquad (8)$$

The value of the SNR for each time-frequency point is obtained as per (8), where the variance of the tuning curve is σ_s^2, the variance of trial-by-trial fluctuations is σ_n^2 and σ_b^2 is a bias correction term introduced to compensate for the low number of trials per direction. If the variance of trial-by-trial fluctuations for a given direction is σ_d, then σ_b is quantified as $\left(\sum_{d=1 \text{ to } 4} \sigma_d^2\right) / (4 * T)$, where T is the number of trials in each target direction. The significant values of SNR are calculated by a randomization test which calculates the significance levels by random shuffling and mislabeling the data and repeating the process 200 times. The significant SNR is thus determined for all the 35 sensors in all seven subjects and the median over all the subjects is reported.
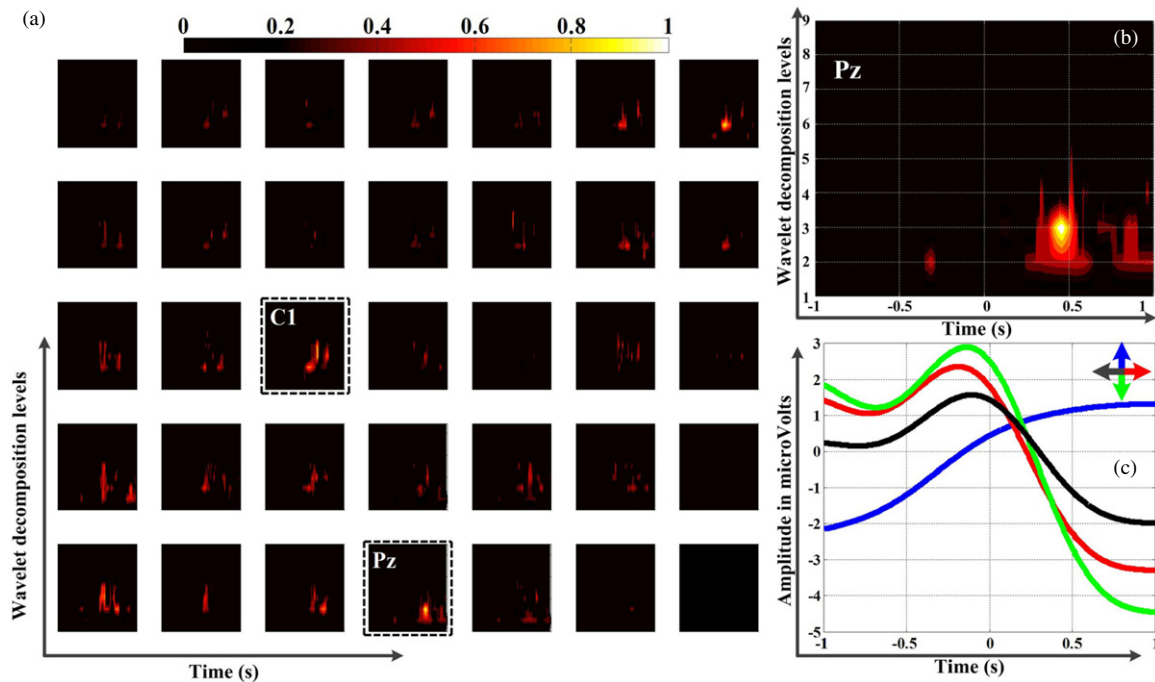
Figure 4. Results of the movement direction dependent SNR study. (a) 35 sensor time-frequency contour plots obtained using wavelet decomposition displaying significant ($p < 0.005$) movement dependent SNR normalized over 35 sensors. The SNR shown is the median over all the subjects. The *y*-axis represents wavelet decomposition levels from 1 to 9 and the *x*-axis provides the time information. These are indicated with gridlines in (b). The labels of sensors with higher significant SNR recorded are shown, Pz, C1. (b) The plot for Pz in (a) is enlarged and the details are shown. It clearly explains the low frequency and movement end activity. (c) The trial-averaged temporal activity for electrode Pz low pass filtered at 1 Hz is shown. The color code for directions is also indicated.

3. Results

This study presents a multi-class feature extraction and classification strategy aiming to analyze EEG data recorded during hand movement in four orthogonal directions. Movement direction dependent amplitude modulation is demonstrated with the aid of time-frequency SNR plots and trial-averaged low frequency EEG time series plots for various hand movement directions. The spatial patterns used in feature extraction and the sensorimotor areas involved in movement activity are also demonstrated with the help of W-CSP patterns and SNR plots. The classification results obtained along with the results of various supporting analyses are provided in this section.

3.1. Significant movement direction dependent SNR

The results of analysis using movement direction dependent SNR, explained in section 2.4 are given in this section. It is essential to identify the temporal frequency regions with high movement direction dependent SNR in order to select the informative levels of wavelet decomposition. The contour plots in figure 4 show the normalized and significant time and frequency bins ($p < 0.005$) in each of the 35 sensors. The figure plots the normalized values of median SNR across all the seven subjects for each sensor. As seen, the electrode C1 in the contralateral motor area and Pz in the midline parietal region provide the highest SNR values compared to others. The SNR information from Pz is plotted separately in figure 4(b). This figure distinctly shows the higher SNR activity towards the movement end at 0.5s and in the decomposition levels less

than 5. As per (1), the frequency regions spanned by the lower five levels are 0.05–0.375, 0.375–0.75, 0.75–1.5, 1.5–3 and 3–6 Hz respectively. In figure 4(c), the trial-averaged 1 Hz low pass filtered signal for each direction recorded from Pz in subject 1 is shown. This shows the amplitude modulation associated with different directions. The results demonstrated in figure 4 justify the selection of the lower five levels of wavelet decomposition in our feature extraction technique. Thus this analysis assists the wavelet approach for extracting features with optimum SNR information.

3.2. Low frequency temporal activity

Many studies [7, 10] have reported the amplitude modulations of neural signals in various frequency bands as being responsible for directional movement. In this section, the sensors that recorded high SNR for movement direction are used to demonstrate the temporal amplitude modulations for each movement direction. The trial-averages of 1 Hz filtered EEG signal in sensors Pz and C1 are shown in figure 5 for each of the seven subjects. The difference in amplitude can be seen in almost all the cases towards the movement end, i.e. after 0.5s and during the early planning period. Also, it can be noted that the temporal trend for direction dependent amplitude modulation is different for different subjects.

3.3. Multi-class classification performance

Figures 6 and 7 summarize the results of cross-validation analysis using various classification algorithms applied on
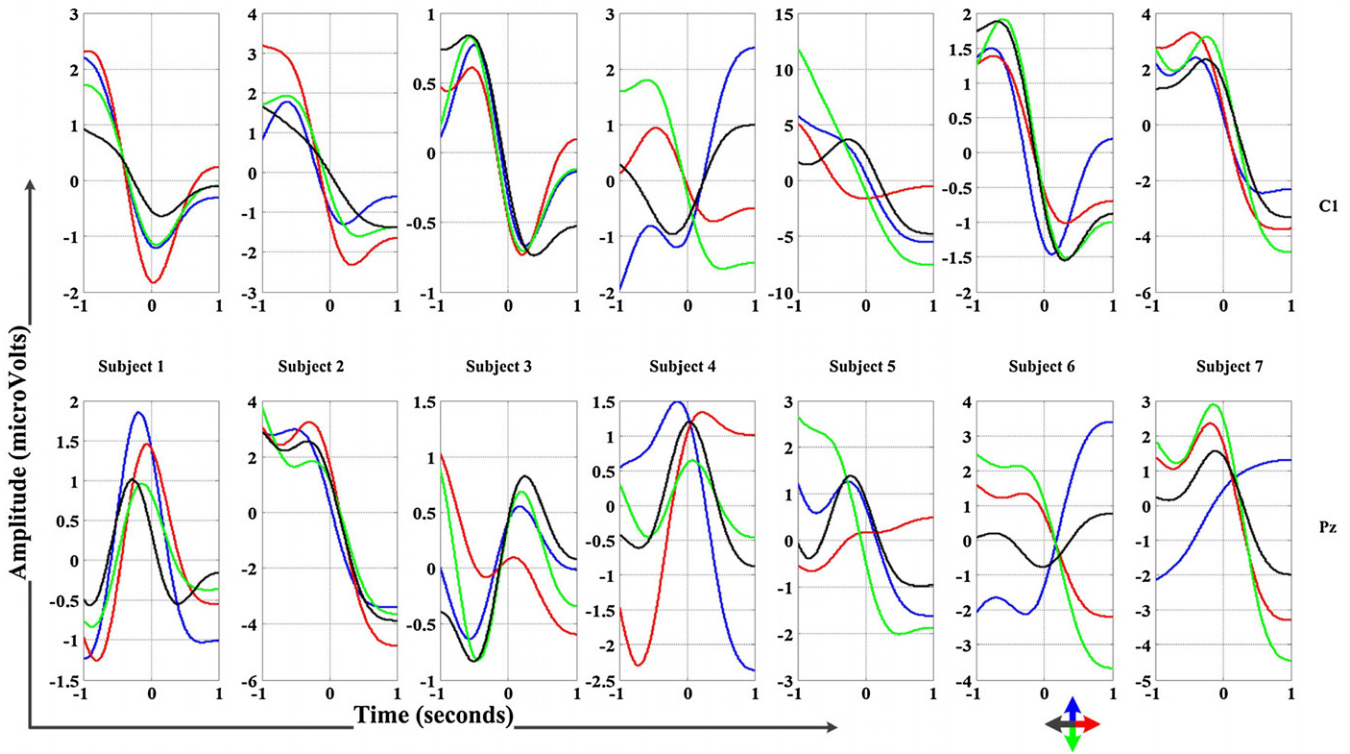
Figure 5. The time series plots of trial-averaged low pass filtered at 1 Hz activity from sensors Pz and C1 are shown. The columns indicate the results for each of the subjects. The color code for directions is also indicated. The distinct amplitude modulations depending on direction can be noticed.
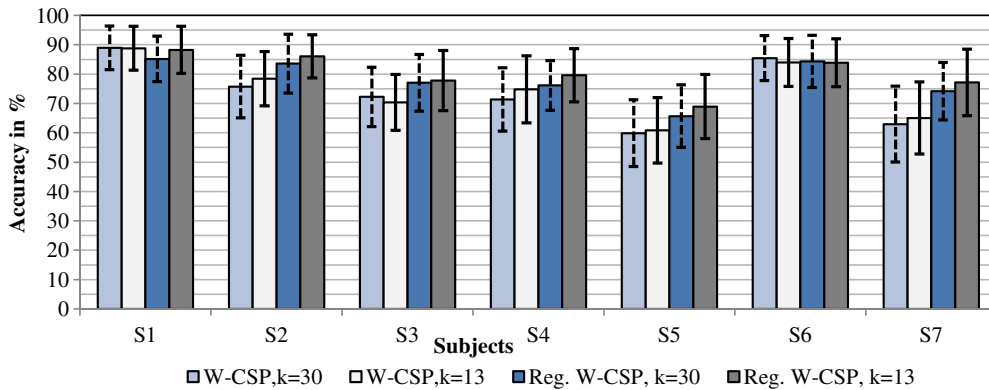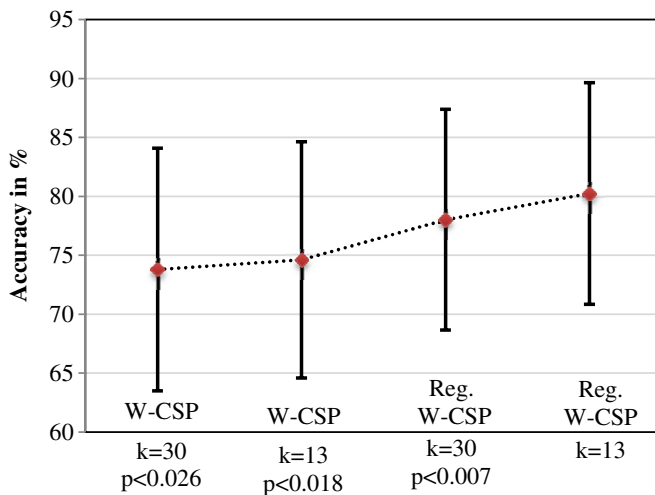

Figure 6. 10 × 10 cross validation accuracies for multi-class classification of hand movement directions. The results for subjects S1 to S7 are given. A comparative performance study using normal [14] and Reg. W-CSP for feature extraction along with and without feature selection are provided. The results conclude that higher performance is obtained in most subjects using Reg. W-CSP followed by FLD classification of selected features.

the data. Regarding feature extraction, the Reg. W-CSP used in this study is compared against normal W-CSP reported in [14] for binary data classification. The figures also shows the results of using the algorithm with and without feature selection. Assuming k to be the number of features used for the classification, the maximum number of features available, i.e. if no feature selection is performed, is $k = 2mL = 30$ ($m = 3$). We select $L = 5$, according to the results in sections 3.1 and 3.2, so as to extract features for the low frequency ($\leqslant 6$ Hz) region. The results obtained by incorporating various feature extraction strategies for individual subjects are shown in figure 6. Figure 7 shows the mean values of classification accuracies along with the

paired t-test results. The number of features selected is set to $k = 13$ for comparisons and the effect of varying k is addressed in figure 8. As shown in figure 7, the proposed method Reg. W-CSP with feature selection gives higher and statistically significant results than normal Reg. W-CSP ($p < 0.007$), normal W-CSP [14] ($p < 0.026$) and W-CSP with feature selection ($p < 0.018$). The Reg. W-CSP provides 4.22% and 5.63% improvement over normal W-CSP for FLD classification without and with feature selection respectively. Feature selection using MI is found to enhance the performance of the Reg. W-CSP algorithm by 2.23%. The improved result clearly shows the use of filtering out the less informative features to provide

Figure 7. Average multi-class classification of hand movement direction performance in terms of accuracy across subjects. A comparative performance study using normal [14] and Reg. W-CSP for feature extraction with and without feature selection is provided. The results conclude that higher and statistically more significant performance is obtained in most subjects using Reg. W-CSP followed by FLD classification of selected features. The results are of paired *t*-test with Reg. W-CSP for $k = 13$.

better classifications. In both algorithms, the classification performance improves by including feature selection. However the increase in W-CSP from $(73.79 \pm 10.29)\%$ to $(74.61 \pm 10.02)\%$ is not statistically significant ($p < 0.334$). In the case of Reg. W-CSP, the increase from $(78.01 \pm 9.37)\%$ to $(80.24 \pm 9.41)\%$ is highly significant ($p < 0.007$). Hence the MI-based feature selection works better in identifying Reg. W-CSP features than in W-CSP features.

The effect of varying the number of features selected, k is summarized in figure 8. The results are for data analysis using Reg. W-CSP, followed by feature selection and FLD classification. On an average over all subjects, the classification performance increases steadily from $k = 1$ to 6 and remains almost constant until $k = 13$ and decreases after $k = 23$. A maximum mean classification accuracy of $(80.24 \pm 9.41)\%$ is obtained for $k = 13$. However, the optimal number of features to be chosen is subject and problem specific. For instance, the number of features providing the best classification performance is 13, 13, 17, 9, 19, 21 and 16 for subjects 1 to 7 respectively.

Figure 9 shows the combined histogram of selected $k = 13$ features for the seven subjects. The horizontal axis gives the feature indices 1 to 30 of which 1–6, 7–12, 13–18, 19–24, 25–30 correspond to six features each from the five sub-bands (3–6, 1.5–3, 0.75–1.5, 0.375–0.75 and 0.05–0.375 Hz respectively). The histogram bars for each class are stacked. A higher value indicates that the feature is selected in most of the subjects in different movement direction classes. For instance, as seen in figure 9, feature 1 is selected from the data of only one of the seven subjects for North, South and West classes (not necessarily the same subject for different classes), whereas feature 29, is selected from all seven subjects for all classes. From the plot, it is evident that for almost all direction classes, the features are mostly selected from the

lowest subband used 0.05–0.375 Hz, thus proving the presence of movement parameter information in low frequency EEG.

3.4. Spatial patterns from Reg. W-CSP

The spatial patterns refer to the projection of brain signal sources to the various EEG sensors obtained using the feature extraction algorithm. Here the spatial patterns from the Reg. W-CSP filters is given by $(W^{l^{-1}})^T$, where W^l is the spatial filter generated at subband *l*. Figure 10 shows the spatial patterns obtained for subject 1. The rows and columns correspond to each of the four direction classes (for the OVR approach) and each of the $L = 5$ subbands used respectively. Being an OVR approach, localized areas defining activity for each movement direction are not well-defined. However, in all the subbands, distinguishable activity for each movement class can be identified.

4. Discussion

The previous section showed the multi-class classification results obtained in our analysis. We further demonstrated the significantly higher direction dependent SNR of low frequency EEG using a SNR study. The results prove the applicability of our algorithm in identifying movement direction information.

4.1. Movement direction encoded in brain recordings: focus on low frequency EEG

As mentioned in section 1, various invasive and non-invasive studies have been used to study center-out multi-directional hand movements. The invasive SUA and MUA approaches used single and multiple neuron firing rates to relate them to movement directions [3, 5, 6]. Many studies have been performed in primates and it was found that the firing rate is cosine-related with movement direction, and this can be used to identify movement directions on a single trial basis. Voluntary control on external effectors was made possible using this approach with MUA from monkey primary motor cortex. Invasive LFPs were also used in a monkey study and this reported low frequency direction tuning and showed its applicability to decode single trial movement direction [7, 8].

Among the non-invasive techniques, a combined study using MEG and EEG used 3 Hz low pass filtered MEG to obtain a decoding accuracy of 67% in a four-direction center-out hand movement experiment. By using EEG alone, 55% decoding accuracy was reported in [3, 10]. The 3 Hz filtered MEG was reported to show different temporal activity for different movement directions. In our current study, a similar activity is found in EEG signals filtered at 1 Hz as shown in figure 5. In this study, our major focus is to identify the most informative low frequency features of EEG. Our results indicate a similar or even better performance using EEG in movement direction classification considering the advantages of EEG over MEG.

The involvement of the parietal region for movement parameter control in non-human primates and humans was reported in various studies [3, 4]. Similarly, activation in
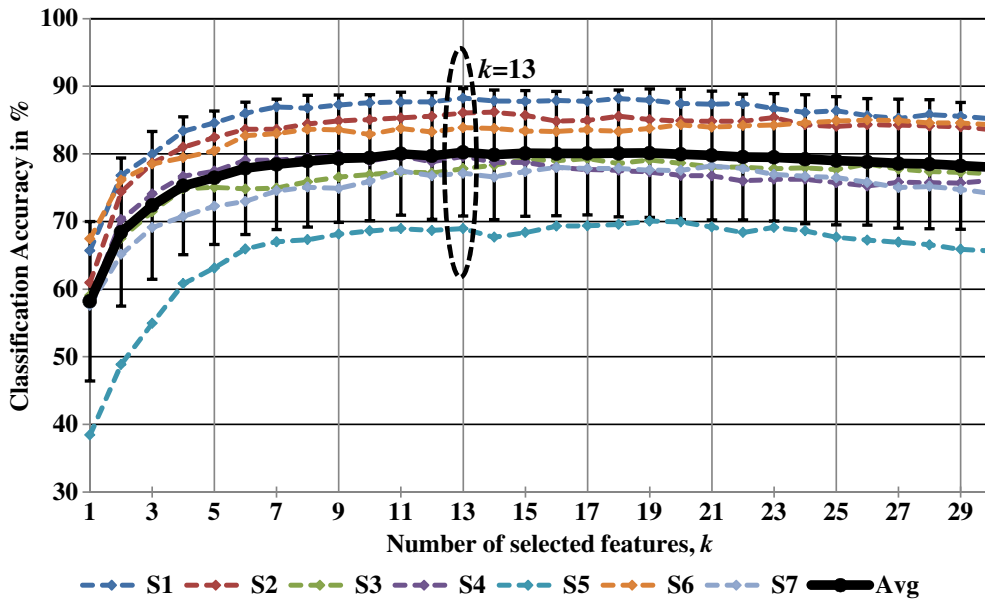
Figure 8. The multi-class classification accuracy as a function of number of features used by the FLD classifier. The legend explains the curves indicating individual subject results and the average over subjects. The average curve indicates an optimal system performance at $k = 13$. However, subject-specific optimal k values are different and are reported in section 3.3.
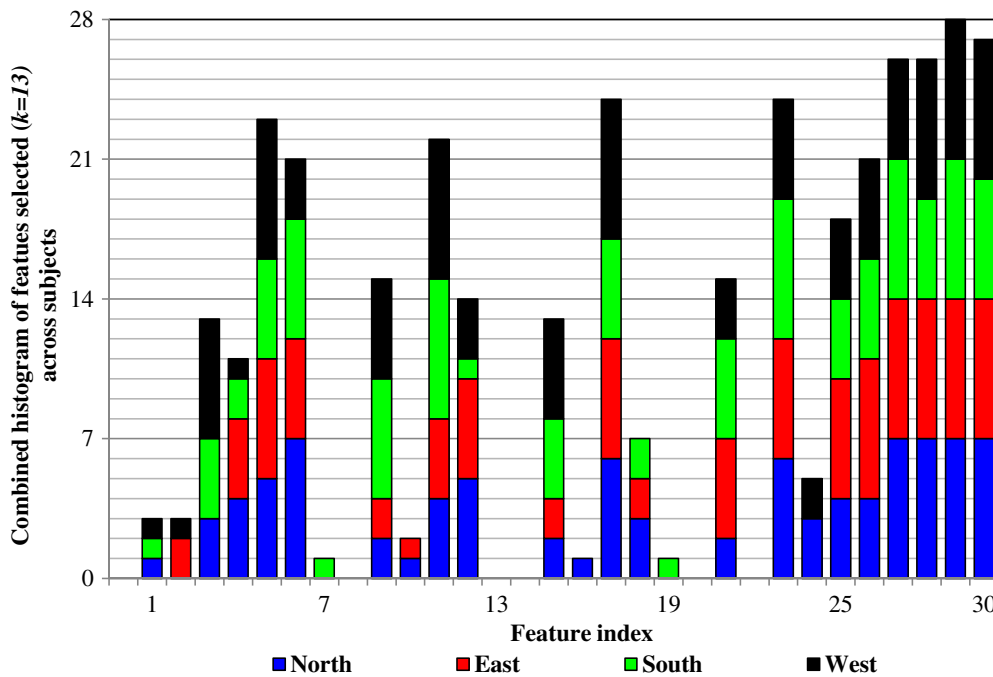

Figure 9. Combined histogram of features selected ($k = 13$) across seven subjects in each movement class. The feature indices corresponding to different frequency subbands are explained in section 3.3. The plot indicates the number of subjects for whom a particular feature is selected from the each movement class data. If a particular feature is selected from all seven subjects for all four movement classes, then combined histogram value is $7 \times 4 = 28$.

contralateral areas is expected for a hand movement task. The role of low frequency brain features in controlling the movement parameters were proved by studies using MEG, ECoG, LFP, etc, in both humans and other primates. The results shown in figure 4 conform to these findings in the literature. The parameter we used is movement direction dependent SNR in each of the temporal-frequency bins and the results show higher a SNR in sensors Pz and C1 for frequencies $\leqslant 6$ Hz.

This strengthens the findings in the literature using various data acquisition modalities.

4.2. Reg. W-CSP approach for movement parameter classification

Another major contribution in this study is the Reg. W-CSP approach that can efficiently extract low frequency information in EEG. A similar approach was used in our previous study
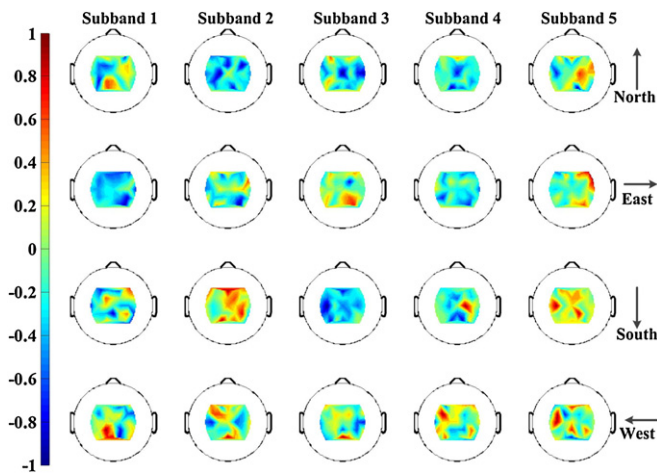
Figure 10. The normalized spatial patterns obtained from Reg. W-CSP spatial filters are displayed in the figure. The results for subject 1 are shown. The OVR approach used designs four filters corresponding to each direction in each of the subbands. The distinct activity in each of the subbands aiding discrimination between various directions can be noted.

[14] in the binary classification of hand movement directions. In the current study, using the proposed approach of the Reg. W-CSP algorithm, followed by MIBIF-based feature selection and FLD classification, a maximum multi-class classification accuracy of $(80.24 \pm 9.41)\%$ is obtained and is found to be more statistically significant than all other methods. Each of these steps filters out the optimal information for the problem at hand. Figure 4 explains the information retrieval using wavelet decomposition. In all the 35 sensors used, the significant movement direction dependent information is obtained only in the lower decomposition levels. The algorithm that we use selects exactly these informative levels to extract features. Figure 8 explains the significance of using MI-based feature selection. As the curve of mean accuracies show, we obtained superior performance $(80.24 \pm 9.41)\%$ by using the selected 13 features rather than using all the features $(78.01 \pm 9.37)\%$. Also, the figure explains the varying trend for different subjects, making this a subject specific problem. The performance of FLD classification is also reported as cross-validation results in figures 6 and 7. Improvement in accuracy is a very important focus of BCI research [2], where algorithms that provide statistically significant and neurophysiologically plausible results would be helpful in building more robust BCI systems [22]. To this end, our proposed approach aims to make a contribution towards developing such accurate and robust BCI systems.

4.3. Experimental limitations

The current study encourages the applicability of non-invasive data acquisition methods in extracting information regarding the movement related parameter, direction. In this study we perform an offline analysis of the recorded EEG data using the proposed feature extraction and classification approach. The results assure us that our strategy can better serve the BCI purpose of refined movement control. However, a practical BCI application aims to create real-time closed loop control

with an output device. In this study we identified the movement direction with high classification accuracy with the help of an offline cross-validation approach. In the feature extraction technique, there are various parameters that can be further optimized. The performance of our algorithm is found to be subject dependent demanding training sessions for all the users and also optimization to select subject specific information. Similarly, the algorithm parameters, such as type of wavelets used, number of features selected, etc, are also problem and user specific. These factors play an important role in setting up a real-time BCI system and need to be addressed in the future.

Furthermore, considering the major application of BCI in the rehabilitation of stroke patients, the demand for identifying movement parameters during motor imagery is important. In this study we focus on actual movement, which is significantly correlated with the neural substrates of motor imagery. In future, we hope to apply our current results and data processing strategy to design online experiments that can continuously classify, decode and reconstruct the trajectory of imagined motor movement.

References

[1] Dornhege G, Millán J R, Hinterberger T, McFarland D J and Müller K R 2007 *Towards Brain–Computer Interfacing* (Cambridge, MA: MIT Press)

[2] Wolpaw J R, Birbaumer N, McFarland D J, Pfurtscheller G and Vaughan T M 2002 Brain-computer interfaces for communication and control *Clin. Neurophysiol.* **113** 767–91

[3] Waldert S, Pistohl T, Braun C, Ball T, Aertsen A and Mehring C 2009 A review on directional information in neural signals for brain–machine interfaces *J. Physiol. Paris* **103** 244–54

[4] Jerbi K, Vidal J R, Mattout J, Maby E and Lecaignard F 2011 Inferring hand movement kinematics from MEG, EEG and intracranial EEG: from brain-machine interfaces to motor rehabilitation *IRBM* **32** 8–18

[5] Georgopoulos A P, Kettner R E and Schwartz A B 1988 Primate motor cortex and free arm movements to visual targets in three-dimensional space: II. Coding of the direction of movement by a neuronal population *J. Neurosci.* **8** 2928–37

[6] Taylor D M, Tillery S I and Schwartz A B 2002 Direct cortical control of 3D neuroprosthetic devices *Science* **296** 1829–32

[7] Rickert J, Oliveira S C, Vaadia E, Aertsen A, Rotter S and Mehring C 2005 Encoding of movement direction in different frequency ranges of motor cortical local field potentials *J. Neurosci.* **25** 8815–24

[8] Mehring C, Rickert J, Vaadia E, Cardosa de Oliveira S, Aertsen A and Rotter S 2003 Inference of hand movements from local field potentials in monkey motor cortex *Nat. Neurosci.* **6** 1253–4

[9] Ball T, Schulze-Bonhage A, Aertsen A and Mehring C 2009 Differential representation of arm movement direction in relation to cortical anatomy and function *J. Neural Eng.* **6** 016006

[10] Waldert S, Preissl H, Demandt E, Braun C and Birbaumer N 2008 Hand movement direction decoded from MEG and EEG *J. Neurosci.* **28** 1000–8

[11] Pfurtscheller G, Neuper C, Flotzinger D and Pregenzer M 1997 EEG-based discrimination between imagination of right and left hand movement *Electroencephalogr. Clin. Neurophysiol.* **103** 642–51

[12] Ball T, Demandt E, Mutschler I, Neitzel E and Mehring C
2008 Movement related activity in the high gamma range of
the human EEG *NeuroImage* **41** 302–10

[13] Demandt E, Mehring C, Vogt K, Schulze-Bonhage A,
Aertsen A and Ball T 2012 Reaching movement onset- and
end-related characteristics of EEG spectral power
modulations *Front. Neurosci.* **6** 65

[14] Robinson N, Vinod A P, Cuntai G, Ang K K and Tee K P 2012
A modified wavelet-common spatial pattern method for
decoding hand movement directions in brain computer
interfaces *Proc. IEEE Int. Joint Conf. on Neural Networks
(Brisbane, June 2012)* pp 2308–12

[15] Krebs H I, Hogan N, Aisen M L and Volpe B T 1998
Robot-aided neurorehabilitation *IEEE Trans. Rehabil. Eng.*
6 75–87

[16] Jung T P, Makeig S, Humphries C, Lee T W and
McKeown M J 2000 Removing electroencephalographic
artifacts by blind source separation *Psychophysiology*
37 163–78

[17] Qin L and He B 2005 A wavelet-based time-frequency
analysis approach for classification of motor imagery for
brain-computer interface applications *J. Neural Eng.*
2 65–72

[18] Yang B H, Yan G Z, Yan R G and Wu T 2007 Adaptive
subject-based feature extraction in brain-computer

interfaces using wavelet packet best basis decomposition
Med. Eng. Phys. **29** 48–53

[19] Mousavi E A, Maller J J, Fitzgerald P B and Lithgow B J 2011
Wavelet common spatial pattern in asynchronous offline
brain computer interfaces *Biomed. Signal Process. Control*
6 121–8

[20] Vetterli M and Herley C 1992 Wavelets and filter banks:
theory and design *IEEE Trans. Signal Process.*
40 2207–32

[21] Ramoser H, Muller-Gerking J and Pfurtscheller G 2000
Optimal spatial filtering of single trial EEG during
imagined hand movement *IEEE Trans. Rehabil. Eng.*
8 441–6

[22] Lotte F and Guan C 2011 Regularizing common spatial
patterns to improve BCI designs: unified theory and new
algorithms *IEEE Trans. Biomed. Eng.* **58** 355–62

[23] Ang K K, Chin Z Y, Wang C, Guan C and Zhang H 2012
Filter bank common spatial pattern algorithm on BCI
competition IV datasets 2a and 2b *Front. Neurosci.*
6 39

[24] Duda R O, Hart P E and Stork D G 2000 *Pattern Classification*
2nd edn (New York: Wiley-Interscience)

[25] McFarland D J, McCane L M, David S V and Wolpaw J R
1997 Spatial filter selection for EEG-based communication
Electroencephalogr. Clin. Neurophysiol. **103** 386–94