PAPER

Wavlet phase-locking based binary classification of hand movement directions from EEG

IOP Publishing Journal of Neural Engineering
J. Neural Eng. **15** (2018) 066008 (13pp) https://doi.org/10.1088/1741-2552/aadeed

Wavlet phase-locking based binary classification of hand movement directions from EEG

Tushar Chouhan[1], Neethu Robinson[1], A P Vinod[2], Kai Keng Ang[1,3] and Cuntai Guan[1]

[1] School of Computer Science and Engineering, Nanyang Technological University, Singapore
[2] Indian Institute of Technology, Palakkad, India
[3] Institute for Infocomm Research, Agency of Science, Technology, and Research (A*STAR), Singapore

E-mail: tushar006@e.ntu.edu.sg

Received 1 November 2017, revised 3 September 2018
Accepted for publication 5 September 2018
Published 25 September 2018

Abstract

Objective. Brain signals can be used to extract relevant features to decode various limb movement parameters such as the direction of upper limb movements. Amplitude based feature extraction techniques have been used to study such motor activity of upper limbs whereas phase synchrony, used to estimate functional relationship between signals, has rarely been used to study single hand movements in different directions. *Approach.* In this paper, a novel phase-locking-based feature extraction method, called wavelet phase-locking value (W-PLV) is proposed to analyse synchronous EEG channel-pairs and classify hand movement directions. EEG data collected from seven subjects performing right hand movements in four orthogonal directions in the horizontal plane is used for this analysis. *Main results.* Our proposed W-PLV based method achieves a mean binary classification accuracy of 76.85% over seven subjects using wavelet levels corresponding to $\leqslant$12 Hz EEG. The results also show direction-dependent information in various wavelet levels and indicate the presence of relevant information in slow cortical potentials ($<$1 Hz) as well as higher wavelet levels ($\leqslant$12 Hz). *Significance.* This study presents a thorough analysis of the phase-locking patterns extracted from EEG corresponding to hand movements in different directions using W-PLV across various wavelet levels and verifies their discriminative ability in the single trial binary classification of hand movement directions.

Keywords: brain–computer interface, phase-locking value, wavelets, electroencephalography, hand movement directions

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

1. Introduction

Brain–computer interfaces (BCI) have been used for several decades to study movement execution and imagination (motor imagery) [1–5]. It is an important area of neuroscience research whose applications include rehabilitation of patients afflicted with neural or spinal cord diseases, partial or full paralysis and stroke [1, 6–8]. Thus, one can find immense motivation to study and develop BCI systems because of its large scope in real world applications and humanitarian technology.

During movement execution and motor imagery, discriminative neurophysiological patterns can be extracted from the recorded brain signals in order to analyse or classify them. The brain coordinates amongst its different specialised regions to achieve a functionally meaningful task [9–11]. This makes it possible to observe different types of phenomena through the recorded brain signals such as local field potentials (LFP), sensorimotor rhythms (SMR), event related potentials (ERP) and visually evoked potentials (VEP) [12].

Amplitude and phase are two important properties of any signal. There are a number of amplitude based features that have been used in literature for decoding the parameters related to limb kinematics. For the purpose of studying arm movements and motor imagery from EEG, the most frequently used methods rely on an amplitude-based spatial filtering technique called common spatial pattern (CSP) [13]. After CSPs were first used in EEG, many variants of CSP have been reported in literature [14–17]. For the purpose of classifying and analysing hand movements in four orthogonal directions in a 2D plane, spatially regularized wavelet common spatial patterns have been shown to perform well with a multi-class classification accuracy of about 80% [18] and binary-class classification accuracy of 71.25% [17].

Like their amplitude-based counterparts, a number of phase-based feature extraction methods have been reported in literature [9, 10, 19, 20]. These include phase lag index (PLI) [20], phase-locking value (PLV) [9, 10, 21], phase slope index (PSI) [22] and imaginary part of coherency (iCOH) [23, 24]. These phase-based features are typically used to study synchrony between different parts of the brain or functional connectivity [9–11]. In other studies, they have been used as feature extraction methods to classify motor imagery of different limbs and actual movement of hands versus rest [21, 25, 26]. Moreover, some studies also show that a combination of amplitude and phase-based features perform better than either type of features leveraging on the fact that both methods essentially estimate different types of neural activity and their respective influence on the measured brain signals [21, 27, 28]. During actual movement execution, amplitude-based CSPs and their numerous variants have been shown to estimate event-related desynchronization/synchronization (ERD/ERS) very well in the mu and beta sensorimotor rhythms [13, 14, 18]. PLVs have been shown to provide important time-domain information for identifying different types of movements such as self-paced or cued movements and have been used to identify channel-pairs that seem to be in task-dependent synchrony [29]. PLVs have also been used to study sensorimotor rhythms during motor imagery and to control BCIs [30–32].

In this paper, we study right hand movements performed in a 2D horizontal plane in 4 orthogonal directions. Previous studies on this and similar data have used variants of CSPs, more specifically wavelet CSP (WCSP) to classify the movement directions [16–18]. However, to the best of our knowledge, a phase-locking based study has not yet been reported in literature to study movement directions. Therefore, this study provides an in-depth analysis of phase-locking based features extracted from the EEG sensor space during right-hand movements in different directions. We also propose an algorithm for binary classification of different direction-pairs and discuss possible ideas for future work.

The remainder of this paper is as follows: section 2 describes the dataset and preprocessing technique used for this study. It also describes the proposed algorithm for binary classification based on wavelet phase-locking values (W-PLV). Section 3 presents the results of binary classification of our proposed method and detailed analyses of the W-PLV features. Section 4 presents discussions on the results, explores ideas for future work and concludes this study.

2. Materials and methods

The purpose of our study is to identify discriminative patterns using phase-locking based features in the EEG sensor space in order to carry out binary classification of hand movement directions using a wavelet filter-bank. Various visualization schemes help us to investigate the wavelet levels and temporal variation in synchrony between channels. We analyse the results of each stage in our proposed algorithm and try to interpret them for possible neurophysiological significance.

2.1. Data acquisition

The experiments were performed at the Neural Signal Processing Lab of Institute of Infocomm Research, Agency for Science, Technology and Research, Singapore. EEG was recorded using a Neuroscan SynAmps 128 channel amplifier of which the 118 EEG channels were used to record the data sampled at 250 Hz. Out of the remaining ten auxiliary channels, two channels were used to record electrooculogram (EOG). Data was collected from seven subjects (ages 25–36 years, mean age 31 years $\pm$ 4.4 [SD], right-handed, all male) performing right hand centre-out movements in the horizontal plane in four orthogonal directions, namely 'north', 'east', 'south' and 'west' denoted by $d1$, $d2$, $d3$ and $d4$, respectively. Each subject was seated on a chair with his right forearm attached to the MIT Manus robot which recorded kinematic parameters such as direction and trajectory of movements [33]. A computer screen displayed distinct cues for movement direction and onset of movement. The experimental timeline has been illustrated in figure 1. The trial began at the home screen which lasted for about 2 s followed by a rest period of 3 s during which the subjects held the MIT MANUS robot in the centre position. The cue for getting ready for hand movement was displayed at the end of the rest period followed by the cue for the direction in which movement is to be performed which signalled the start of the preparation interval. As it was a randomized controlled experiment, in each trial the direction prompted was chosen randomly by a computer. The preparation interval which lasted for 2 s was followed by a cue for onset of movement. This cue for movement onset was identical for all directions. Subjects were instructed to carry out the movement as soon as they received the cue for onset of movement and to complete the movement within 0.5 s. At the end of the task, the subject was informed that the trial was complete. A total of 160 trials per subject were conducted (40 trials per direction) for six subjects, S1–S6 and 140 trials (35 trials per direction) for subject S7. The total time for each trial took around 10 s. For this study, a subset of 35 EEG electrodes spanning the sensorimotor cortex, as shown in figure 1, and two EOG electrodes are used.

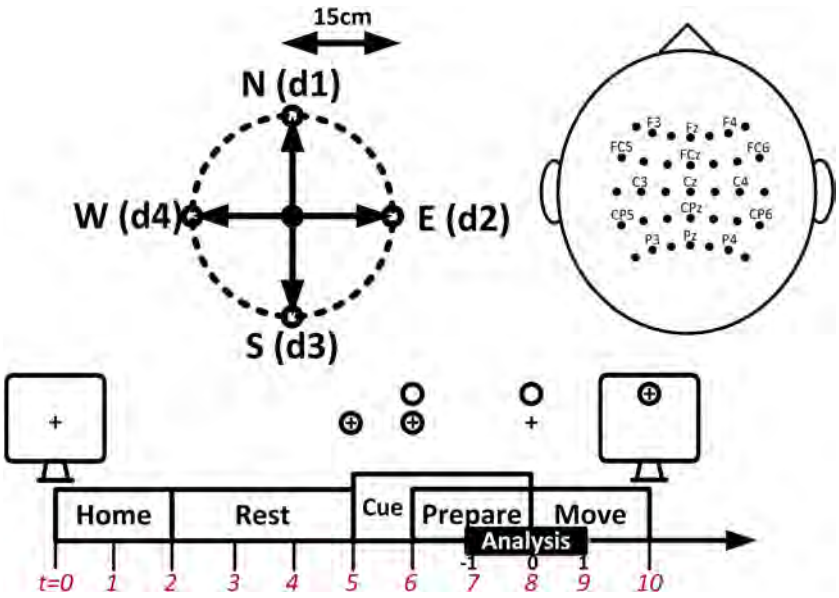

Figure 1. Experimental timeline of each task and the set of 35 EEG electrodes that have been used for this study which include F5, F3, F1, Fz, F2, F4, F6, FC5, FC3, FC1, FCz, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6, P5, P3, P1, Pz, P2, P4 and P6. The symbols above the timeline show the various cues, relative positions on the screen and the time instants at which they were displayed to the subjects during each trial.

2.2. Preprocessing of EEG data

The acquired signals have to be preprocessed to enhance their quality and remove artefacts due to eye movements (EOG) and muscle activity known as electromyogram (EMG). In our study, 2 s of data are considered for each trial, which correspond to the interval between 1 s prior to movement onset cue to 1 s after it, as shown in figure 1. Thus, there are 500 samples of EEG data corresponding to 2 s per trial. All data processing is done in MATLAB. Firstly, the acquired signals are low pass filtered at 96 Hz using a zero phase Chebyshev Type II filter. The filter order is determined from the data using MATLAB's cheby2ord(·) function which calculates the minimum order required to meet the filter design specifications. The cut off frequency is 96 Hz with less than 3 dB of passband ripple and 40 dB of stopband attenuation. Notch filtering is performed to remove line frequency at 50 Hz using a second order IIR notch filter (iirnotch(·) in MATLAB). Since the data acquisition hardware had a lower cut-off frequency of 0.05 Hz, the spectrum of EEG signals used in this study is between 0.05–96 Hz. Independent component analysis (ICA) is then carried out on the signals for EOG artefact removal based on correlation with the EOG signals [34]. The components with highest correlation with the EOG signals are nullified. The conditioned EEG signals are then back-projected to the sensor space. Surface Laplacian is then used to remove the higher frequency artefacts due to EMG [35]. Apart from EOG and EMG artefacts, we did not find any evidence of electromagnetic field interference in the EEG data. More details of the above pre-processing techniques can be found in [18].

2.3. Proposed methodology for binary classification

The method used for single-trial binary classification of hand movement directions in this paper is similar to our earlier work [36] wherein we use Symlets5 wavelet to successively decompose and reconstruct the preprocessed EEG signals at various frequency bands. These narrowband signals are then used for extracting wavelet phase-locking values (W-PLV) as features for single-trial binary classification. However, unlike our previous work, we use a simplified version of the correlation-based channel-pair selection technique (feature selection) and include higher wavelet levels corresponding to frequency $\leqslant$12 Hz which includes the mu band that is typically used to study movement and motor imagery parameters, besides the slow cortical potentials (SCPs) [37, 38]. We use 5×5 nested cross-validation to assess the performance of our binary classification algorithm. Feature selection is only done on the training set which is then applied to the test set. The results of the classification performance reported are then the average accuracy on the test set over all five runs of 5-fold cross-validation. We also investigate the contribution of different wavelet levels to the classification performance achieved and attempt to find evidence of information through various trial-averaged methods.

2.3.1. Discrete wavelet transform and wavelet reconstruction. Wavelets offer an effective way of analysing EEG signals since they offer excellent frequency resolution in the lower frequencies [39, 40]. The prototype functions are dilated and scaled to shift across time and frequency which are then used as basis functions for analysing and reconstructing the EEG signals [39, 41]. For digital signal processing, instead of performing continuous wavelet transform, we perform discrete wavelet transform by discretizing the translation and dilation parameters as this helps to remove redundancy caused by continuous parameters. This discretization is done on a dyadic grid to get the resultant spectrum division as illustrated in [41]. In this study, 500 samples of data are divided into nine

wavelet levels. The frequency bands corresponding to the nine wavelet levels are: 0.05–0.375, 0.375–0.75, 0.75–1.5, 1.5–3, 3–6, 6–12, 12–24, 24–48 and 48–96 Hz. We consider wavelet levels from 1 to 6 (1 being the wavelet level corresponding to the lowest frequency band) and ignore wavelet levels 7–9. This is because the higher wavelet levels (7–9) are broadband EEG frequency bands and phase-locking values are usually computed for narrow EEG frequency bands. Moreover, wavelets 1–6 already include SCPs and the mu band which are typically used to study motor activity.

2.3.2. Wavelet-phase locking value. PLV is a measure of phase synchronization between a pair of signals and is computed using the phase-difference between the pair of signals [11]. A constant phase-difference would imply perfectly synchronized signals and would result in a high PLV of 1. If the signals vary randomly with respect to each other then the PLV is 0. We use PLV to identify pairs of EEG channels that show synchronous relationships for different hand movement directions. PLVs can be computed from the phase of a narrowband signal computed using the Hilbert transform as defined in equations (1) and (2) [19].

$$x_h(t) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(t)}{t-\tau}d\tau \tag{1}$$

$$\phi_x(t) = arg(x(t) + jx_h(t)) \tag{2}$$

where $x_h(t)$ is the result of the Hilbert transform of the narrowband signal $x(t)$ that is obtained by computing the integral taken in the sense of Cauchy principal value, $arg(\cdot)$ is the argument function that computes the instantaneous phase of the signal, denoted by $\phi_x(t)$.

Once we compute the instantaneous phase of the signal, the trial-averaged PLV between two channels x and y are computed using equations (3) and (4).

$$\epsilon(x,y,t) = \exp j(|\phi_x(t) - \phi_y(t)|) \tag{3}$$

$$V_{d,l}(x,y,t) = \left|\frac{1}{T_d}\sum_{trials\in d}\epsilon(x,y,t)\right| \tag{4}$$

where $x(t)$ and $y(t)$ are the EEG signals from channels x and y at wavelet level l, $t \in [1,N]$ where N is the number of time samples (500) in each trial, $d \in D$ where D is the set of all directions $= \{d1, d2, d3, d4\}$ as shown in figure 1. T_d is the number of trials corresponding to direction d. $\epsilon(x,y,t)$ is a measure of the single trial phase-locking between the two channels at time t and $V_{d,l}(x,y,t)$ is the PLV calculated as the average over all the trials belonging to that direction. Since we compute the PLV values after performing wavelet reconstruction of the signals, we call them Wavelet-PLV or W-PLV.

W-PLV values computed above are then used to compute W-PLS values, with the help of a statistical randomization test as follows. At every reconstructed wavelet level and for a given pair of electrodes, a surrogate W-PLV distribution is generated for every time instant. This is done by randomly labelling and shuffling the data for one of the channels in every channel pair and re-computing the W-PLV values. This process is repeated ten times the sample size, T_d for each subject to generate the surrogate distribution. Then the original W-PLV values are compared against this surrogate distribution using the Student's t-test to calculate W-PLS, using a similar technique as in [9]. As shown in equation (5), we check for a significance level of $\alpha = 0.001$ and accept the W-PLV values if they are found to be significant, else the corresponding W-PLS values are taken as 0. For 5-fold cross validation, the sample size T_d includes the training data wherein there are 32 trials per direction for subjects S1–S6 and 28 trials per direction for S7.

$$S_{d,l}(x,y,t) = \begin{cases} V_{d,l}(x,y,t) & \text{if } p\text{-Value} \leqslant \alpha \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $S_{d,l}(x, y, t)$ is the W-PLS value for direction class d and wavelet level l. In order to carry out single trial classification, we use single trial W-PLVs as features from all combinations of channel-pairs. Thus, given a channel pair (x, y), we calculate the single trial PLV by using equation (6).

$$\epsilon_{k,l}(x,y) = \left|\frac{1}{N}\sum_{t=1}^{N}\epsilon_k(x,y,t)\right| \tag{6}$$

where $\epsilon_{k,l}(x, y)$ is the single trial W-PLV for the kth trial at wavelet level l.

2.3.3. Pearson's correlation coefficient based channel-pair selection. As mentioned earlier, in this study we use 35 channels which are directly above the motor cortex. There are 595 combinations of channel-pairs for which we can compute W-PLV at every wavelet level. For six wavelet levels, there are a total of 3570 features that can be computed using W-PLV which would lead to a high dimensional feature matrix for single trial classification. Therefore, we use Pearson's correlation coefficient (PCC) to select the best channel-pairs from each wavelet level that show most discriminative information for a given pair of directions (for binary classification). PCC is often used in statistics to test both, the qualitative as well as the quantitative nature of the correlation between two variables [42]. The correlation coefficient 'r' varies between -1 and 1 such that when $r = 1$ the variables have perfect positive correlation, i.e. they increase and decrease together perfectly; and when $r = -1$ the variables have perfect negative correlation, i.e. they have perfectly opposite trends.

We use PCC in our method to select channel-pairs which have the highest PCC values with the class labels, as shown in equations (7) and (8).

$$Z_1 = [V_{di,l}(x,y,t), V_{dj,l}(x,y,t)] \forall t \in [1,N] \tag{7}$$

$$Z_2 = [L_1, L_{-1}] \tag{8}$$

where the vector Z_1 of length $2N$ contains the appended trial-averaged W-PLV values for binary-classes (di, dj) where $(di, dj) \in D$ and $di \neq dj$, for a channel-pair (x, y) at wavelet level l, Z_2 is a vector of length $2N$, whose elements, vectors L_1 and L_{-1}, are each of length N and contain labels 1 and -1 respectively. We compute PCC for all combinations of channel-pairs, at every wavelet level and sort them in descending order

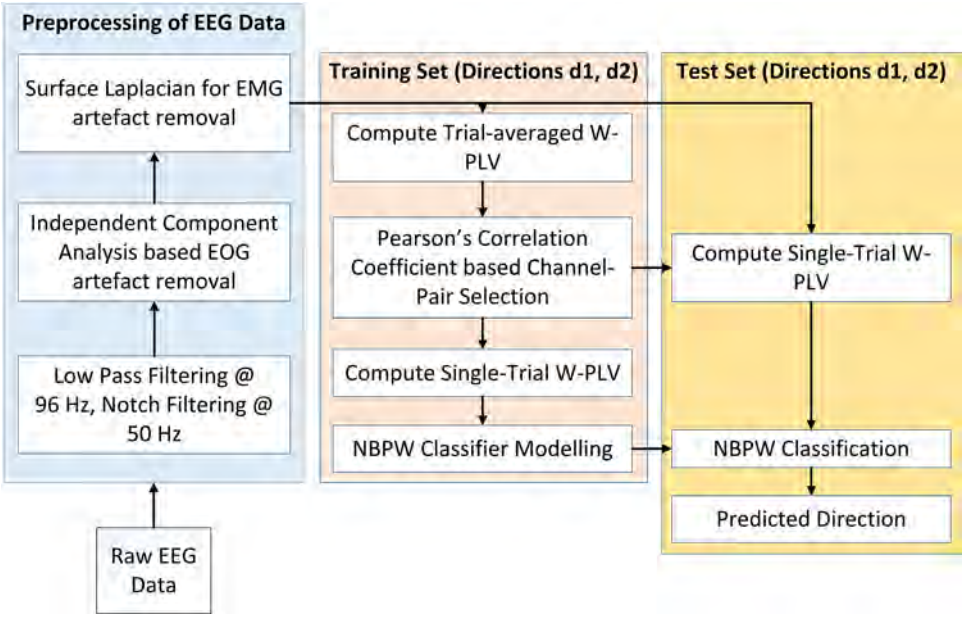

Figure 2. Proposed methodology for binary classification using W-PLV.

from which we select the channel-pairs corresponding to the highest positive and negative PCC values. In this study, we select 10 channel-pairs each corresponding to the correlation values from both ends of the sorted PCC values resulting in 20 channel-pairs at every wavelet level. $\tau_{di,dj,l}$ contains this set of channel-pairs selected for a set of binary-classes (di, dj) at every wavelet level l. The corresponding values of single trial WPLV calculated using (6) are stored in $\mathbf{F}$. This is then the feature matrix that is used for single-trial classification.

2.3.4. Naive Bayesian Parzen window classifier (NBPW). A Naive Bayesian Parzen window classifier is used for binary classification, which employs Bayes rule to predict the class which has maximum posterior probability $p(d|f)$ as computed in equation (9) [15].

$$p(d|f) = \frac{p(f|d)\,p(d)}{p(f)} \tag{9}$$

where $p(d|f)$ is the posterior probability for direction d and feature f wherein f is a feature vector $\in \mathbf{F}$ that corresponds to the W-PLV values computed for a channel-pair selected in $\tau_{d_1,d_2,l}$. For 2 classes, the probability $p(f)$ is calculated as follows.

$$p(f) = \sum_{d=1}^{2} p(f|d)\,p(d). \tag{10}$$

Since it is a naive classifier, it assumes that all the features are independent. On the basis of this assumption, the conditional probabilities are calculated as follows.

$$p(f|d) = \prod_{j=1}^{d} p(f_j|d) \tag{11}$$

$$\hat{p}(f_j|d) = \frac{1}{n_d} \sum_{i \in I_d} \phi(f_j - \bar{f}_{i,j}, h) \tag{12}$$

where $\hat{p}(f_j|d)$ is the Parzen Window estimator of the conditional probability $p(f_j|d)$, n_d is the number of data samples belonging to class d, I_d is the set of indices of training data trials $\in d$, ϕ is a smoothing kernel function with a smoothing parameter as defined in [15].

2.3.5. Algorithm. The steps in the binary classification algorithm, illustrated in figure 2, are summarized in this section as follows. The dataset is pre-processed and segregated into training and test sets for 5×5 nested cross-validation. For a pair of binary classes, di and $dj \in \{d1, d2, d3, d4\}$ and $di \neq dj$, the training data is used to compute the trial averaged W-PLV values. As mentioned earlier, we use only the lower 6 wavelet levels corresponding to the frequency bands 0.05–0.375, 0.375–0.75, 0.75–1.5, 1.5–3, 3–6 and 6–12 Hz respectively. Like the methodology proposed in [36], we select channel-pairs using Pearson's correlation coefficient. However, instead of trial-averaged W-PLS values as used in [36], we use the trial averaged W-PLV values for channel-pair selection. This is because computation of W-PLS is extremely complex and time-consuming as it is heavily iterative. For this study, we use W-PLS values only to visually inspect the data. Since we are using PCC to select the best channel-pairs at every wavelet level, it is expected that the channel-pairs with significant W-PLV values would be picked automatically.

Feature selection is only done using the training data. As we select 20 channel-pairs per wavelet level, there are a total of 120 features selected over all six wavelet levels. These selected channel-pairs are then used to compute the single-trial W-PLV

values from the training set to train the NBPW classifier. After five runs of 5-fold cross-validation, we report the average classification accuracy over the test set. The steps have been summarized in algorithm 1 and illustrated in figure 2.

Algorithm 1. Wavelet-PLV based binary classification.

1: Segregate trials into training and test sets for 5×5 cross-validation
2: **for** each cross-validation fold **do**
3: **for** trials $\in$ training set and binary-class (di, dj) **do**
4: Compute $V_{d,l}(x, y, t) \forall t \in [1, N], d = \{di, dj\}$ for all combinations of (x, y)
5: Perform PCC at every wavelet level l and choose subset of channel-pairs $\tau_{di,dj,l}$
6: **end for**
7: $\mathbf{F}_{train} = v_{k,l}(x, y) \forall$ trials $\in$ training set $\forall (x, y) \in \tau_{di,dj,l}$.
8: Train NBPW Classifier
9: $\mathbf{F}_{test} = v_{k,l}(x, y) \forall$ trials $\in$ test set $\forall (x, y) \in \tau_{di,dj,l}$.
10: Test the performance of the NBPW classifier on test set and calculate classification accuracy
11: **end for**
12: Repeat above steps for remaining iterations of 5×5 cross-validation.
13: Calculate average classification accuracy over test set

3. Analysis and results

This section presents the results of binary classification of hand movement directions with a detailed analysis of trial-averaged and single-trial W-PLV features. In doing so, we attempt to seek greater insight into the movement related spectro-temporal information in the various wavelet levels. Thus, we first analyse the trial-averaged W-PLV features extracted from EEG corresponding to the different hand movement directions, as described in section 2. Next, we present the results of binary classification of all combinations of hand movement directions using single-trial W-PLV features. Lastly, we investigate the results of the different stages of the binary classification algorithm to gain better understanding of the working of the algorithm. We also try to provide possible interpretations of the associated neurological phenomena, wherever possible.

3.1. W-PLV plots

3.1.1. Temporal variation of windowed trial-averaged W-PLS. We would like to visualize the temporal variation of trial-averaged W-PLS values over the 2 s of data per trial we use in this study. The motivation behind such inspection is to investigate the presence of any temporal markers that could give us some movement-related information. For this purpose, we divide the entire duration of 500 samples (2 s) into time-windows of 100 samples each wherein each window would correspond to an interval of 0.4 s. Next, we compute the mean of the trial-averaged W-PLS values computed above over all the samples within that window as in equation (13).

$$\bar{S}_{w_{d,l}}(x, y) = \frac{1}{100} \sum_{t \in w} S_{d,l}(x, y, t) \tag{13}$$

where $\bar{S}_{w_{d,l}}$ is the windowed-trial averaged W-PLS value for a channel pair (x, y) and direction d, at wavelet level l and time-window w. Thus, there are a total of five windows, each of duration 0.4 s.

Figure 3(a) shows the number of statistically significantly phase-locked channel-pairs in each time window (TW1–TW5) for each direction averaged over all subjects for wavelet level 3. It can be seen that for all directions, TW3 and TW4 contain the maximum number of significantly phase-locked channel-pairs than all other TWs. For all directions except direction $d3$, TW3 has maximum number of significantly phase-locked channel-pairs. For direction $d3$, TW4 has maximum number of significantly phase-locked channel-pairs. Figure 3(b) has been plotted similar to those in [30] as an example to visualize the temporal patterns of W-PLS values. For this, we use the imagesc function of MATLAB and plot the channel $\times$ channel matrix for direction $d2$ of subject 1. As the channel $\times$ channel matrices are symmetric, only the upper triangles have been plotted in figure 3 for better comprehension. We can observe the build-up of phase-locking patterns starting from TW1, peaking at TW3 and reducing again in TW5, following the temporal sequence observed for all subjects in figure 3(a). TW3 corresponds to -0.2 to 0.2 s, that includes the cue for the onset of movement at time $t = 0$. In this particular time window, it is observed that channels C6, CP5 and CP2 (corresponding to columns 21, 22 and 26) have maximum number of peaks in W-PLS values with other channels. A possible interpretation of this observation is that the activity in this time window as measured on the scalp through EEG might be centered around C6, CP5 and CP2. As this time window corresponds to the visual cue for onset of movement, the increased synchrony observed on these channels may indicate the increased interaction at the somatosensory cortex. This observation, though expected, will have to be verified using source-imaging techniques. The same channels also show huge temporal variation in the estimated synchrony and depict a build-up of synchrony starting with almost none in the first time window, peaking at the third time window and reducing to negligible synchrony again in the fifth time window. Thus, the number of peaks in the windowed, trial-averaged W-PLS values may help us in isolating period of onset of motor execution. Our future work will use this temporal information for identifying the above.

3.2. Spatial contour maps

W-PLV values averaged and normalized over all subjects are next analysed in order to identify some common trends in phase-locking patterns for the different directions. For this purpose, figure 4 is plotted which presents the spatial contour maps of the mean W-PLV values over all subjects, using EEGLAB in MATLAB [43]. These spatial contour maps are

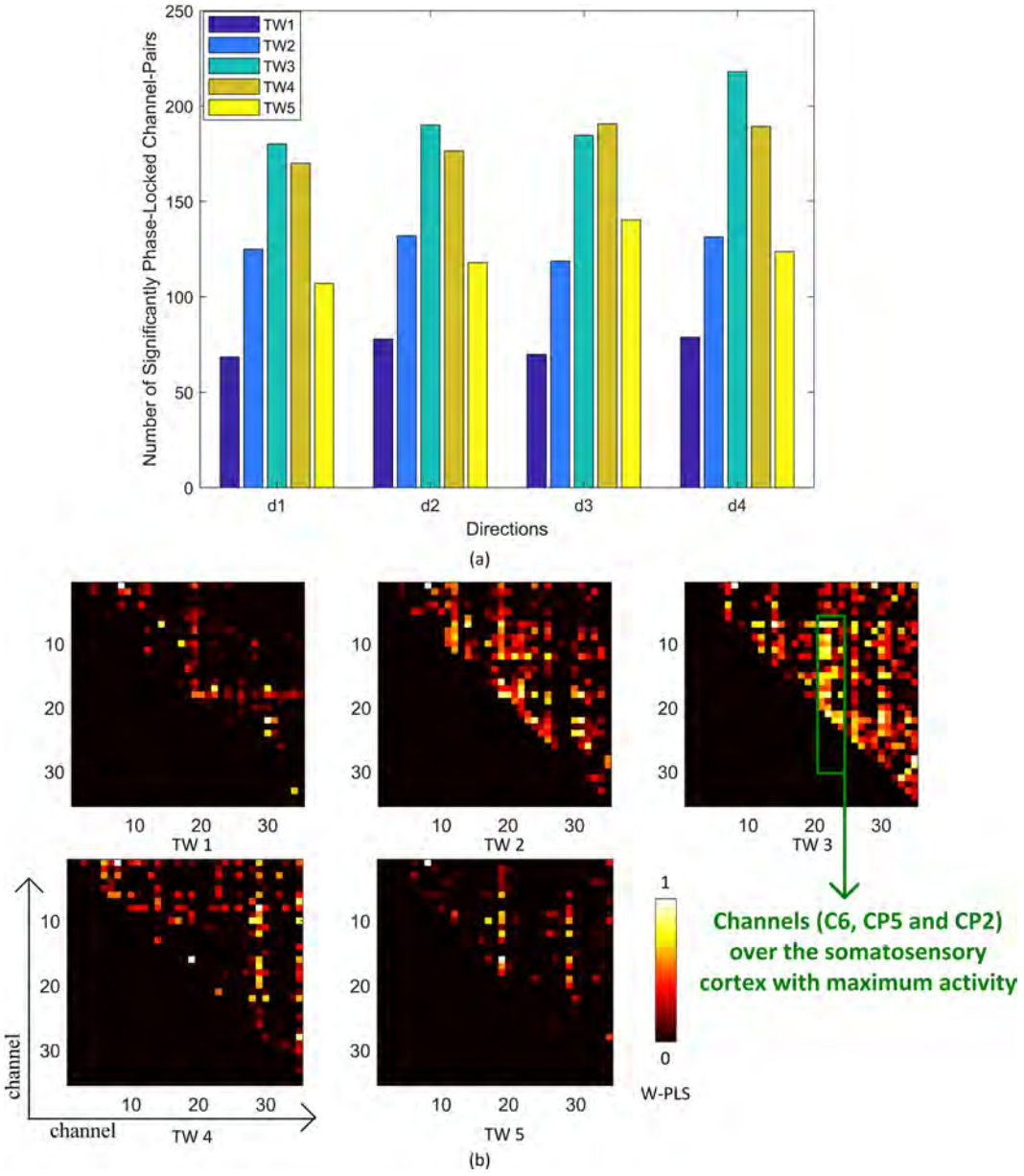

Figure 3. Analysis of the temporal variation of windowed W-PLS values. (a) Plot of the number of significantly phase-locked channel-pairs for each direction in each time window averaged over all subjects for wavelet level 3 (0.375–0.75 Hz). (b) Example of significantly phase-locked channels plotted using direction *d*2, subject S1. The 35 channels in each channel × channel subplot are arranged as F5, F3, F1, Fz, F2, F4, F6, FC5, FC3, FC1, FCz, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6, P5, P3, P1, Pz, P2, P4 and P6 from top to bottom and from left to right (1 to 35, respectively). As each channel × channel matrix is symmetric, only the upper triangle has been plotted for better comprehension. TW refers to the time window. The five time windows (from TW1 to TW5) correspond to time durations −1 to −0.6, −0.6 to −0.2, −0.2 to 0.2, 0.2 to 0.6 and 0.6 to 1 s, respectively. Time window TW3 shows maximum interaction centred around channels C6, CP5 and CP2 which are directly over the somatosensory cortex.

plotted using W-PLV values computed for all channels with respect to the common 'seed' channel Cz. After calculating the mean W-PLV values over all the subjects for each direction and wavelet level, values are then normalized such that the minimum and maximum W-PLV values take 0 and 1 respectively.

From figure 4, it can be seen that the lowest wavelet levels have distinctly most variation in W-PLV patterns as compared to the higher levels. Peaks on these plots imply increased interaction between the respective channels and seed channel Cz. The high variation in the normalized W-PLV patterns seen could imply direction-dependent information extracted from EEG which would help our channel-pair selection technique select the most discriminative channel-pairs as will be seen in section 3.4. In wavelet level 1, denoted by *l*1, the plots for all the 4 directions show common peaks at C2 while the peaks on the channels on the left hemisphere (contralateral to the moving arm) show varying patterns. The opposite trend is observed for wavelet level 2 wherein peaks can be seen on C1 and FC1 (left hemisphere) but varying patterns are seen on the channels of the right hemisphere (ipsilateral to the moving arm). There are differences in the contours for the higher wavelet levels as well but the peaks are most distinctly different for the different directions in the lower wavelet

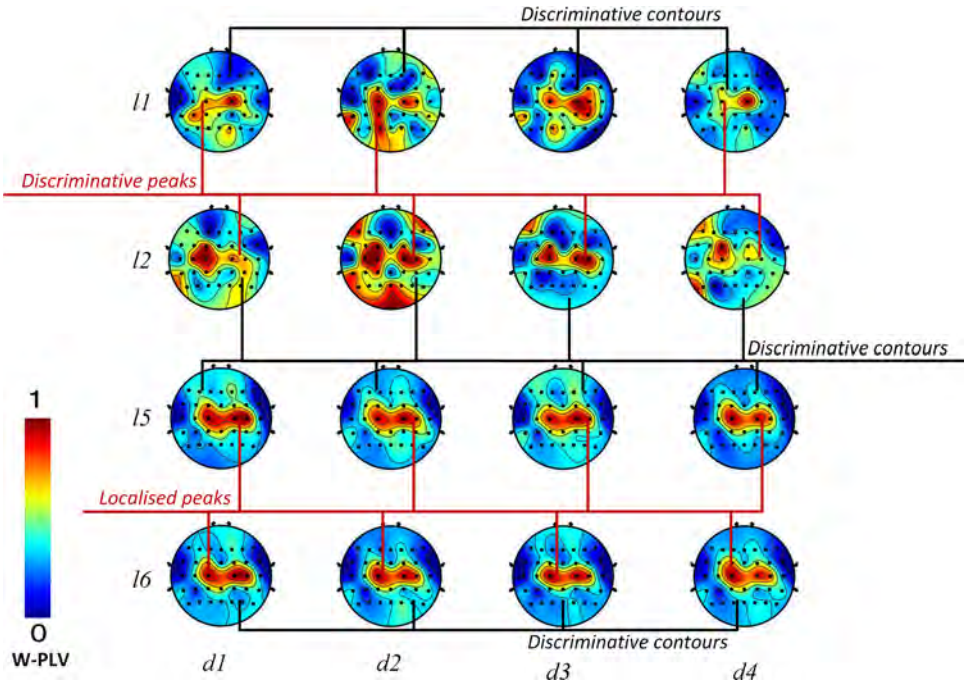

Figure 4. Spatial contour maps plotted using W-PLV values, averaged over all trials and all subjects with seed channel at Cz. Wavelet levels *l*1, *l*2, *l*5 and *l*6 correspond to frequency bands 0.05–0.375 Hz, 0.375–0.75 Hz, 3–6 Hz and 6–12 Hz, respectively. The spatial maps in levels *l*1 and *l*2 seem to have most direction-dependent W-PLV patterns while those in *l*5 and *l*6 seem to have peaks localized around C1 and C2 and direction-dependent contours. The peaks on the left hemisphere (contralateral to the moving arm) in *l*1 and on the right hemisphere (ipsilateral to the moving arm) in level *l*2 seem to be distinctly direction-dependent. Identification of such peaks can help to increase the classification accuracy of the different hand movement directions.

Table 1. Results of binary classification for subjects S1–S7 using our proposed W-PLV method with Nave Bayesian Parzen window classifier after 5 × 5 nested cross-validation. Mean accuracies reported below in percentages.

	Binary classes						
Subjects	*d*1–*d*2	*d*1–*d*3	*d*1–*d*4	*d*2–*d*3	*d*2–*d*4	*d*3–*d*4	Mean
S1	91.75	81.00	89.25	81.25	91.00	82.00	86.04
S2	71.25	68.00	69.00	67.75	78.50	78.25	72.12
S3	69.75	71.25	78.00	75.25	95.25	72.75	77.04
S4	79.00	69.75	79.25	77.75	74.75	89.75	78.37
S5	82.25	64.00	71.25	76.75	74.00	77.25	74.25
S6	76.00	84.25	80.25	76.25	78.75	79.50	79.71
S7	67.43	76.86	72.29	65.71	68.86	74.57	70.95
Mean	76.78	73.59	77.04	74.39	80.16	79.15	76.85

levels. For wavelet levels 5 and 6, as shown in figure 4, the peaks seem to be localised at channels around C1 and C2. On changing the seed channel, different phase-locking patterns and contours are observed. However, irrespective of the choice of seed channel, the contrast in the visually discriminative information in the lower and upper wavelet levels, as shown in figure 4 is consistent for all such figures.

Table 2. Comparison of results of our proposed W-PLV method with other methods that have been used previously on the same dataset. Accuracies are reported below in percentages.

Methods	Mean accuracy
Proposed W-PLV method	76.85
Modified W-CSP [17]	71.25
W-PLV [36]	65.78

3.3. Binary classification results

Single-trial binary classification of hand movement directions is conducted for all six combinations of binary classes of the four hand movement directions. These results are tabulated for all seven subjects using our proposed method in table 1. The mean accuracy achieved after 5 × 5 nested cross-validation over all 7 subjects is 76.85%. The 95% confidence interval for the mean classification accuracy is (72.13%, 81.57%) which is statistically significantly above chance level, as computed using Student's One-Sample *t*-test ($p < 0.00001$). We see empirically that including more wavelet levels increases the performance of binary classification to 76.85% as compared to [36] where only the lowest two wavelet levels have been used and a mean classification accuracy of 65.7% is achieved, as shown in table 2. Moreover, in table 2, the classification performance of our W-PLV method is compared with the

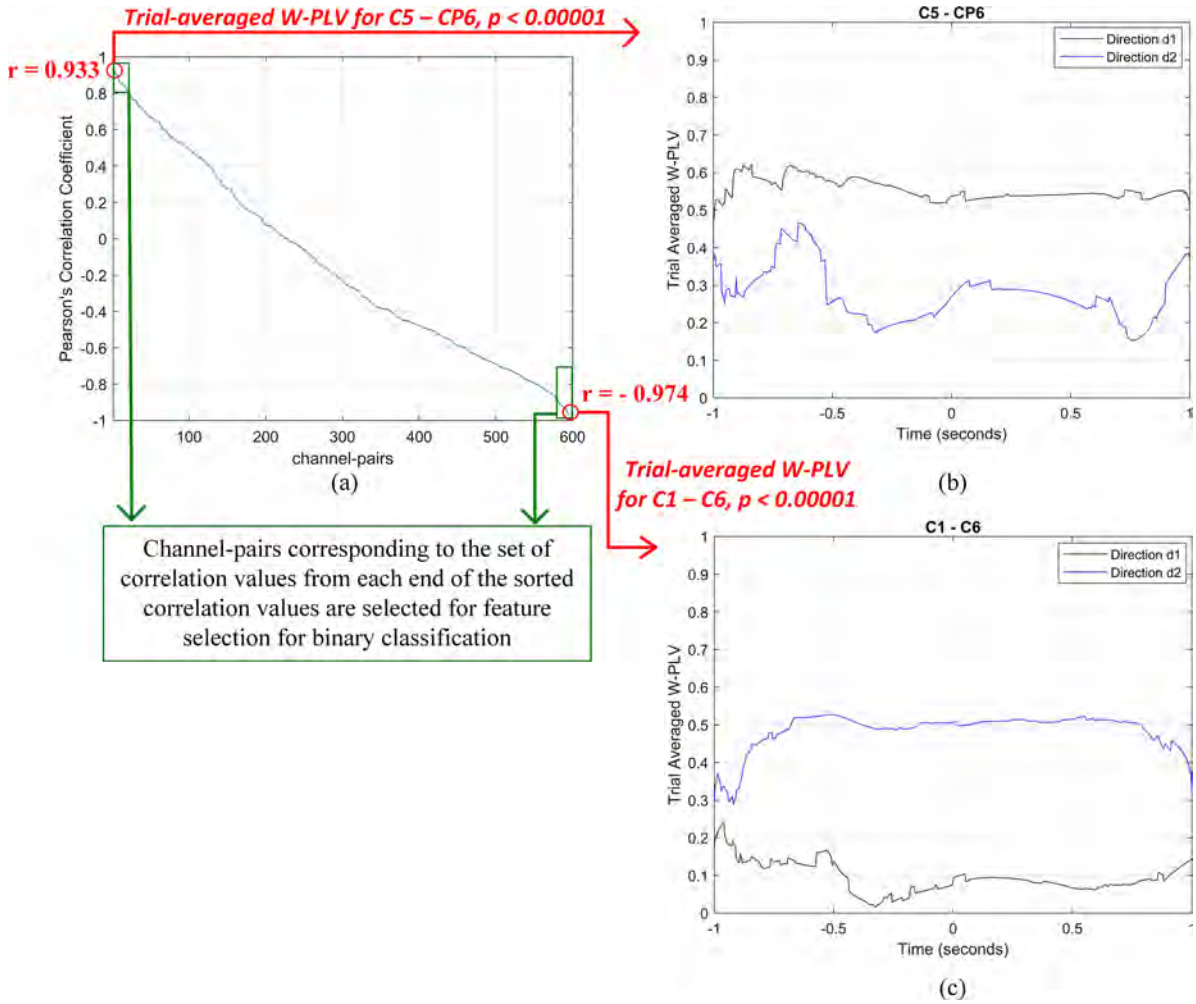

Figure 5. Example of PCC based channel-pair selection on trial-averaged W-PLV values with the correlation coefficients sorted in descending order in (a). This figure has been plotted for subject S1 and wavelet level 1 for the binary class of directions $d1$ and $d2$. The channel-pairs corresponding to the correlation values enclosed in the green boxes on either side of the sorted PCC values show most discriminative information between the two classes and are hence selected for feature selection for binary classification. The plots in (b) and (c) are two examples of the selected channel-pairs at each end of the sorted correlation values, with maximum positive correlation and maximum negative correlation respectively, which are selected for feature selection for binary classification. In these figures, the plots in black correspond to the trial-averaged W-PLV for direction $d1$ and blue to direction $d2$. PCC for each set of signals are mentioned in red for each figure.

state-of-the-art modified wavelet CSP method in [17] for binary classification and is seen to give better results. Next, we used Student's paired t-test to compare the performance of our proposed method with the methods mentioned in table 2. We found that although the mean accuracy of our proposed method is higher than WCSP, this difference is not statistically significant ($p = 0.1891$). The 95% confidence interval of WCSP in [17] is (61.11%, 81.37%). Thus, our proposed method results in a much narrower confidence interval indicating its robustness. When compared to our earlier W-PLV based method in [36], our proposed method performed significantly better ($p = 0.006$).

The trend of classification performance over the seven subjects is different in W-PLV and WCSP [17] methods which are based on phase-synchrony and amplitude of EEG signals, respectively. This could mean that there is complementary information in these different feature extraction methods. An effective combination of these features could help to further increase the classification accuracy.

3.4. Channel-pair selection using Pearson's correlation coefficient

We would like to analyse the working of our modified correlation-based channel-pair selection technique. For this purpose we plot figure 5 for subject S1 for channel-pairs that show highest positive correlation and highest negative correlation for wavelet level 1 and directions $d1$ and $d2$. In this figure, the graph in (a) shows the correlation coefficients sorted in decreasing order for all the 595 channel-pairs with the respect to directions $d1$ and $d2$. The higher positive PCC values correspond to those channel-pairs that show increased phase-locking for direction $d1$ and low phase-locking for direction $d2$. Channel-pair C5-CP6 has the highest positive PCC value of 0.933 and the trial averaged W-PLV values for both the directions have been plotted in figure 5 showing consistently higher values for direction $d1$ as compared to direction $d2$. The opposite trend can be seen for channel-pair C1–C6 which have highest negative correlation with PCC value of -0.9743

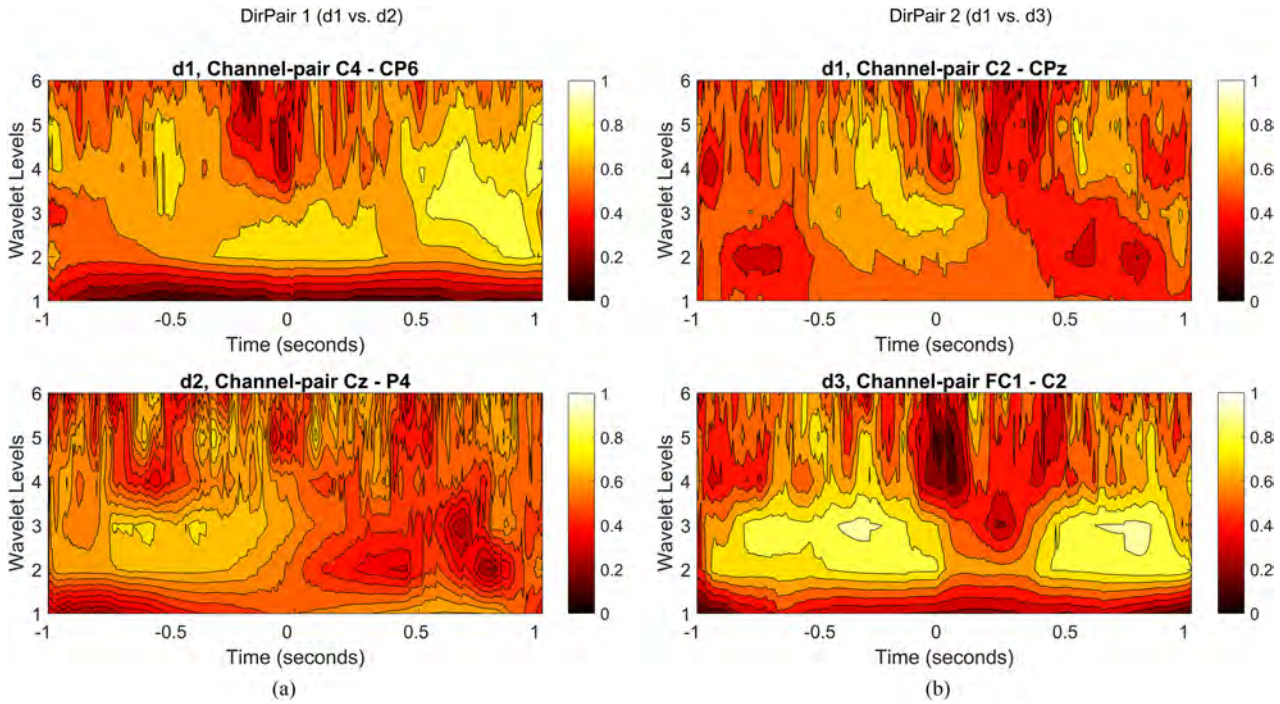

Figure 6. Time-frequency plot of some of the most often selected channel-pairs whose W-PLV values are used for binary classification for subject S1, showing the presence of direction-dependent W-PLV peaks across various wavelet levels. (a) and (b) Have been plotted for features selected most often for binary-classes $d1$ versus $d2$ and $d1$ versus $d3$, respectively. In (a), channel-pair C4-CP6 is chosen most often for $d1$ and Cz-P4 is chosen most often for $d2$ across all wavelet levels. In (b), channel-pair C2-CPz is chosen most often for $d1$ and FC1–C2 is chosen most often for $d3$. Peaks in both (a) and (b) can be seen spread out across all the wavelet levels from 1 to 6. Lower wavelet levels correspond to lower frequency bands.

with the class labels as mentioned in (8). The channel-pairs corresponding to the PCC values in the region enclosed in the green boxes on either side of the sorted PCC values, which include the channel-pairs in (b) and (c) are selected for feature selection for binary classification.

3.5. Time-frequency contour plots of most frequently selected channel-pairs

It is seen that including higher wavelet levels has helped to increase the classification accuracy of our W-PLV based method significantly from 65.7% in [36] to 76.85%. We would now like to investigate the contribution of the higher wavelet levels included to the increase in binary classification accuracy. Our hypothesis is that single-trial features from the higher wavelet levels containing direction-dependent information should contain consistent peaks in trial-averaged W-PLV values. We test this hypothesis by plotting time-frequency contour maps of frequently chosen features chosen across all wavelet levels. Presence and location of such peaks would provide relevant movement related information.

Figure 6 has been plotted using some features that are most often selected for binary classification for subject S1 as an example. From this figure we see that W-PLV peaks can be seen across all wavelet levels. In the lower levels (levels 1–3 corresponding to EEG < 1 Hz), there are peaks in C4-CP6, from time -0.4 to 0.4 s in direction $d1$ and in Cz-P4 in direction $d2$ from time -0.8 to -0.4 s of figure 6(a). In figure 6(b), W-PLV peaks in levels 2–3 can be clearly seen in FC1–C2, direction $d3$. Due to the higher temporal resolution of the lower levels, these peaks appear to be very prominent. Due to the nature of the temporal resolution of wavelets, the peaks in the higher levels are much narrower as compared to those in the lower levels. However, we can still spot them clearly in levels 5 and 6 (corresponding to EEG between 3–12 Hz) in Cz-P4 in figure 6(a) in time -0.5 to -0.2 s and 0 to 0.2 s. In C2-CPz in figure 6(b) W-PLV peaks in the levels 5–6 can be seen in time -0.4 to -0.2 s and 0 to 0.12 s; in FC1–C2 in time -0.6 to -0.4 s and 0 to 0.2 s. The same observation can be made from similar figures plotted for other subjects and direction-pairs. The presence of such direction-dependent peaks in higher wavelet levels explains why including more wavelet levels adds discriminative information to the feature matrix that is used by the classifier and results in higher classification performance. Thus, they seem to have contributed to the increase in classification accuracy. Moreover, even though $d1$ is common in (a) and (b), the features selected most often for $d1$ in both cases are different as they depend on the other class. This is because the feature selected should contain most discriminative information that would help in single trial classification. Hence, our proposed method is capable of intelligently identifying discriminative features for binary classification.

4. Discussion

This paper began with a literature review on the various methods that have been used to extract hand movement-related information. After that, a novel W-PLV-based feature extraction method was proposed to analyse right hand movement in

4 orthogonal directions in a horizontal 2-dimensional plane. The analyses on the proposed feature extraction method utilizing wavelet phase-locking values from EEG demonstrate the ability of the proposed method in identifying task-dependent synchronous channel-pairs. The discriminative ability of our proposed method is verified in the binary classification of hand movement directions which yields higher classification performance than other methods reported in literature.

4.1. Discriminative patterns of hand movement directions in W-PLV

The wavelet phase-locking patterns extracted have the advantage of being visualized effectively which helps reveal discriminative patterns. Spatial contour maps plotted from trial-averaged W-PLV values show that the lowest two wavelet levels (<1 Hz) contain the most movement-related information. However, there are also visible differences in contours in higher wavelet levels as seen in figure 4. Figure 3 depicts the temporal variation of statistically significant phase-locking in various channel-pairs. Inspection of this figure shows the presence of relevant movement related information such as the markedly increased phase-locking in channel-pairs centred over the somatosensory cortex during the period of onset of movement. Thus, it may be possible to use such timing information to enhance the efficiency of a movement detector which can then be used for asynchronous classification using EEG. Studies have shown the ability of using PLV-based techniques to identify actual movement and motor imagery from rest [29, 44]. Moreover, it might be possible to develop a strategy to use such information to fine-tune a detector which identifies the best time-period for classification. It is important to note that identical cues were used for movement onset for all directions and therefore contained no discriminative information.

We have conducted a thorough analysis on the results of various steps used in the proposed W-PLV based binary classification method and empirical evidence is presented to corroborate the improvement in classification performance. Even though the trial-averaged W-PLV patterns in figure 4 are distinctly more discriminative for the lowest wavelet levels than the higher ones, figure 6 shows the presence of movement related peaks in W-PLV values in higher wavelet levels as well. Thus, inclusion of single-trial features extracted from higher wavelet levels adds discriminative information that can be used by the classifier.

4.2. Limitations

The primary limitation of the current work is that the feature extraction method and binary classification algorithm is computationally intensive. Since all possible combinations of channel-pairs are considered, there are a total of 595 channel-pairs between the 35 channels used for this study. Increasing the number of channels used for this study would increase the number of combinations of channel-pairs which would drastically increase the computational complexity of the algorithm. A possible way to mitigate this limitation would be to use a channel selection algorithm to choose a smaller subset of the most relevant channels before computing the W-PLVs between the channel-pairs. Alternatively, the number of wavelet levels used for classification can be fine-tuned according to each subject.

The current method would have to be modified when extended to multi-class classification of hand movement directions. Commonly used methods for multi-class classification such as one-versus rest (OVR) when directly applied to the current method might lead to poorer accuracies due to the small size of the current dataset (40 trials per class per subject for six subjects and 35 trials per class for the seventh subject) and would instead lead to the problem of imbalanced training data, biasing the classification towards the 'rest' class. One would then have to use more sophisticated techniques to treat imbalanced data such as cost-sensitive boosting [45, 46] and resampling [46, 47]. Secondly, the current correlation-based feature selection method is appropriate for binary classes only as it selects features with the highest correlation coefficients with the class labels in positive and negative extremities. Thus, the present method would have to undergo relevant modifications to address the above issues before extending it to multi-class classification. Once this is achieved, we can compare the performance of our modified W-PLV based method against established amplitude-based multi-class classification strategies like the one reported in [18].

4.3. Future work

The amplitude and phase of a signal extract different kinds of information about the recorded activity. Therefore, future efforts in this line of work will be made to explore machine learning methods that effectively combine the different kinds of information from amplitude and phase-based feature extraction techniques. Also, as observed earlier, the difference in trend of classification accuracies of the subjects in W-PLV and WCSP methods imply complementary information in these methods that could be used to further enhance the classification performance. A possible strategy is to use ensemble classifiers that would use different types of features to classify hand-movement directions and then strategically combine the predicted labels of the individual classifiers to make a collective decision.

Acknowledgment

The authors acknowledge the support of the Ministry of Education, Singapore, under the AcRF Tier 2 Grant MOE2015-T2-2-011.

ORCID iDs

Tushar Chouhan https://orcid.org/0000-0001-8090-2079
Neethu Robinson https://orcid.org/0000-0001-8531-2940
Kai Keng Ang https://orcid.org/0000-0002-3053-6311
Cuntai Guan https://orcid.org/0000-0002-0872-3276

References

[1] Daly J J and Wolpaw J R 2008 Brain–computer interfaces in neurological rehabilitation *Lancet Neurol.* **7** 1032–43

[2] Pfurtscheller G and Da Silva F L 1999 Event-related EEG/MEG synchronization and desynchronization: basic principles *Clin. Neurophysiol.* **110** 1842–57

[3] Nicolas-Alonso L F and Gomez-Gil J 2012 Brain computer interfaces, a review *Sensors* **12** 1211–79

[4] 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

[5] Wolpaw J and Wolpaw E W 2012 *Brain–Computer Interfaces: Principles and Practice* (Oxford: Oxford University Press)

[6] Birbaumer N and Cohen L G 2007 Brain–computer interfaces: communication and restoration of movement in paralysis *J. Physiol.* **579** 621–36

[7] Buch E *et al* 2008 Think to move: a neuromagnetic brain–computer interface (BCI) system for chronic stroke *Stroke* **39** 910–7

[8] Daly J J, Cheng R, Rogers J, Litinas K, Hrovat K and Dohring M 2009 Feasibility of a new application of noninvasive brain computer interface (BCI): a case study of training for recovery of volitional motor control after stroke *J. Neurol. Phys. Therapy* **33** 203–11

[9] Lachaux J P, Rodriguez E, Martinerie J and Varela F J 1999 Measuring phase synchrony in brain signals *Hum. Brain Mapp.* **8** 194–208

[10] Lachaux J P, Rodriguez E, Le Van Quyen M, Lutz A, Martinerie J and Varela F J 2000 Studying single-trials of phase synchronous activity in the brain *Int. J. Bifurcation Chaos* **10** 2429–39

[11] Bastos A M and Schoffelen J M 2015 A tutorial review of functional connectivity analysis methods and their interpretational pitfalls *Frontiers Syst. Neurosci.* **9** 175

[12] Van De Velde M 2000 Signal validation in electroencephalography research *PhD Thesis* Technische Universiteit Eindhoven (https://doi.org/10.6100/IR529756)

[13] 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

[14] 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

[15] Ang K K, Chin Z Y, Zhang H and Guan C 2008 Filter bank common spatial pattern (FBCSP) in brain–computer interface *Proc. IEEE Int. Joint Conf. on Neural Networks (Hong Kong, China, June 2008)* pp 2390–7

[16] Robinson N, Vinod A P, Guan C, Ang K K and Peng T K 2011 A wavelet-CSP method to classify hand movement directions in EEG based BCI system *Proc. Int. Conf. on Information, Communications and Signal Processing (Singapore, December 2011)* pp 1–5

[17] Robinson N, Vinod A P, Guan C, Ang K K and Peng T K 2012 A modified wavelet-common spatial pattern method for decoding hand movement directions in brain computer interfaces *Proc. IEEE Int. Joint Conference on Neural Networks (Brisbane, Australia, June 2012)* pp 1–5

[18] Robinson N, Guan C, Vinod A P, Ang K K and Peng T K 2013 Multi-class EEG classification of voluntary hand movement directions *J. Neural Eng.* **10** 056018

[19] Le Van Quyen M, Foucher J, Lachaux J, Rodriguez E, Lutz A, Martinerie J and Varela F J 2001 Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony *J. Neurosci. Methods* **111** 83–98

[20] Vinck M, Oostenveld R, Van Wingerden M, Battaglia F and Pennartz C M 2011 An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias *NeuroImage* **55** 1548–65

[21] Hamner B, Leeb R, Tavella M and Millán J D R 2011 Phase-based features for motor imagery brain–computer interfaces *Annual Int. Conf. IEEE Eng. Med. Biol. Soc. (Boston, USA, August 2011)* (IEEE) pp 2578–81

[22] Thatcher R W, Palmero-Soler E, North D M and Biver C J 2016 Intelligence and EEG measures of information flow: efficiency and homeostatic neuroplasticity *Sci. Rep.* **6** 38890

[23] Borich M R, Wheaton L A, Brodie S M, Lakhani B and Boyd L A 2016 Evaluating interhemispheric cortical responses to transcranial magnetic stimulation in chronic stroke: a TMS-EEG investigation *Neurosci. Lett.* **618** 25–30

[24] Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S and Hallett M 2004 Identifying true brain interaction from EEG data using the imaginary part of coherency *Clin. Neurophysiol.* **115** 2292–307

[25] Jian W, Chen M and McFarland D J 2017 EEG based zero-phase phase-locking value (plv) and effects of spatial filtering during actual movement *Brain Res. Bull.* **130** 156–64

[26] Townsend G and Feng Y 2008 Using phase information to reveal the nature of event-related desynchronization *Biomed. Signal Process. Control* **3** 192–202

[27] Hsu W Y 2015 Enhancing the performance of motor imagery eeg classification using phase features *Clin. EEG Neurosci.* **46** 113–8

[28] He W, Wei P, Zhou Y and Wang L 2012 Combination of amplitude and phase features under a uniform framework with EMD in EEG-based brain–computer interface *Annual Int. Conf. IEEE Eng. Med. Biol. Soc. (San Diego, USA, August 2012)* pp 1687–90

[29] Popovych S, Rosjat N, Toth T I, Wang B A, Liu L, Abdollahi R O, Viswanathan S, Grefkes C, Fink G R and Daun S 2016 Movement-related phase locking in the delta-theta frequency band *NeuroImage* **139** 439–49

[30] Wang Y, Hong B, Gao X and Gao S 2007 Design of electrode layout for motor imagery based brain–computer interface *Electron. Lett.* **43** 557–8

[31] Brunner C, Scherer R, Graimann B, Supp G and Pfurtscheller G 2006 Online control of a brain–computer interface using phase synchronization *IEEE Trans. Biomed. Eng.* **53** 2501–6

[32] Jian W, Chen M and McFarland D J 2017 Use of phase-locking value in sensorimotor rhythm-based brain–computer interface: zero-phase coupling and effects of spatial filters *Med. Biol. Eng. Comput.* **55** 1915–26

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

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

[35] 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

[36] Chouhan T, Robinson N, Vinod A P and Ang K K 2017 Binary classification of hand movement directions from eeg using wavelet phase-locking *Proc. IEEE Int. Conf. on Systems, Man and Cybernetics (Banff, Canada, October 2017)* pp 264–9

[37] McFarland D J, Miner L A, Vaughan T M and Wolpaw J R 2000 Mu and beta rhythm topographies during motor imagery and actual movements *Brain Topography* **12** 177–86

[38] Pfurtscheller G and Neuper C 2001 Motor imagery and direct brain–computer communication *Proc. IEEE* **89** 1123–34
[39] Vetterli M and Herley C 1992 Wavelets and filter banks: theory and design *IEEE Trans. Signal Process.* **40** 2207–32
[40] Mallat S 1999 *A Wavelet Tour of Signal Processing* (New York: Academic)
[41] Robinson N, Vinod A P, Ang K K, Tee K P and Guan C T 2013 EEG-based classification of fast and slow hand movements using wavelet-CSP algorithm *IEEE Trans. Biomed. Eng.* **60** 2123–32
[42] Benesty J, Chen J, Huang Y and Cohen I 2009 Pearson correlation coefficient *Noise Reduction in Speech Processing* (Berlin: Springer) pp 1–4
[43] Delorme A and Makeig S 2004 EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis *J. Neurosci. Methods* **134** 9–21
[44] Sweeney-Reed C M and Nasuto S J 2009 Detection of neural correlates of self-paced motor activity using empirical mode decomposition phase locking analysis *J. Neurosci. Methods* **184** 54–70
[45] Sun Y, Kamel M S, Wong A K and Wang Y 2007 Cost-sensitive boosting for classification of imbalanced data *Pattern Recognit.* **40** 3358–78
[46] Sun Y, Wong A K and Kamel M S 2009 Classification of imbalanced data: a review *Int. J. Pattern Recognit. Artif. Intell.* **23** 687–719
[47] Drosou K, Georgiou S, Koukouvinos C and Stylianou S 2014 Support vector machines classification on class imbalanced data: a case study with real medical data *J. Data Sci.* **12** 143–55 (http://www.jds-online.com/files/JDS-1310.pdf)