IOP Publishing Journal of Neural Engineering
J. Neural Eng. **11** (2014) 035016 (13pp) doi:10.1088/1741-2560/11/3/035016

Detection of motor imagery of swallow EEG signals based on the dual-tree complex wavelet transform and adaptive model selection

Huijuan Yang[1], **Cuntai Guan**[1], **Karen Sui Geok Chua**[2], **See San Chok**[2], **Chuan Chu Wang**[1], **Phua Kok Soon**[1], **Christina Ka Yin Tang**[1] **and Kai Keng Ang**[1]

[1] Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore
[2] Tan Tock Seng Hospital Rehabilitation Centre, Singapore

E-mail: hjyang@i2r.a-star.edu.sg

Received 20 March 2014, revised 26 March 2014
Accepted for publication 26 March 2014
Published 19 May 2014

Abstract

Objective. Detection of motor imagery of hand/arm has been extensively studied for stroke rehabilitation. This paper firstly investigates the detection of motor imagery of swallow (MI-SW) and motor imagery of tongue protrusion (MI-Ton) in an attempt to find a novel solution for post-stroke dysphagia rehabilitation. Detection of MI-SW from a simple yet relevant modality such as MI-Ton is then investigated, motivated by the similarity in activation patterns between tongue movements and swallowing and there being fewer movement artifacts in performing tongue movements compared to swallowing. *Approach.* Novel features were extracted based on the coefficients of the dual-tree complex wavelet transform to build multiple training models for detecting MI-SW. The session-to-session classification accuracy was boosted by adaptively selecting the training model to maximize the ratio of between-classes distances versus within-class distances, using features of training and evaluation data. *Main results.* Our proposed method yielded averaged cross-validation (CV) classification accuracies of 70.89% and 73.79% for MI-SW and MI-Ton for ten healthy subjects, which are significantly better than the results from existing methods. In addition, averaged CV accuracies of 66.40% and 70.24% for MI-SW and MI-Ton were obtained for one stroke patient, demonstrating the detectability of MI-SW and MI-Ton from the idle state. Furthermore, averaged session-to-session classification accuracies of 72.08% and 70% were achieved for ten healthy subjects and one stroke patient using the MI-Ton model. *Significance.* These results and the subjectwise strong correlations in classification accuracies between MI-SW and MI-Ton demonstrated the feasibility of detecting MI-SW from MI-Ton models.

Keywords: motor imagery, brain–computer interface, swallowing, dysphagia, tongue protrusion

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

1. Introduction

Dysphagia is the inability to swallow or difficulty in swallowing caused by stroke or other neuro-degenerative diseases. Post-stroke dysphagia has been documented in approximately 30–42% of acute stroke patients requiring hospital admission [1]. Problems in post-stroke dysphagia range from oral transit phase issues to reduced oropharyngeal control and post-pharyngeal phase issues. Conventional treatment methods include: dietary changes, food and

position changes, tongue strengthening exercises, pharyngeal maneuvers, and thermal and neuro-muscular stimulation methods such as VitalStim [2]. Recently, transcranial magnetic stimulation has been used to reorganize the human motor cortex, on the basis of the frequency, intensity and duration of the stimulus [3]. Control of swallowing generally occurs at two levels of brainstem and cerebral cortex [4]. Swallowing is a complex process involving sensory processing of ingested materials, oral–pharyngeal–laryngeal motor coordination, integration with respiration and mastication components, and finally overall cognitive and attentional processing [5, 6].

Evidence that motor imagery activates similar pathways to executed movements [7] has motivated the use of motor imagery for the rehabilitation of stroke patients [8, 9]. Motor imagery, i.e., mental rehearsal or simulating the physical movements, appears to provide an alternative way to access the motor system for rehabilitation. Mental practice has been used in combination with actual practice to rehabilitate motor deficits in patients with subacute stroke. Improvements have been shown in strength, function, and the use of both upper and lower extremities in chronic stroke [10]. The increase in stride length after performing the motor imagery exercise therapy yielded improvements in gait speed [11]. While motor imagery may be a useful tool in activating hand and limb areas in the brain [7–9], it has never been used as a potential therapeutic strategy for swallowing. Shared brain activation areas of swallowing and tongue elevation including the left lateral pericentral and anterior parietal cortex, and anterior cingulated cortex and adjacent supplement motor area suggest the possibility of using motor imagery of tongue protrusion as a means of detecting the motor imagery of swallow [6]. In addition, Furlong *et al* suggested that sensory input to the tongue, tongue thrust and wet swallow activated similar swallowing-related areas in the brain such as the posterolateral sensorimotor and primary gustatory cortex, both critical in the central control of swallowing [12]. Accurate detection of the motor imagery of swallow (MI-SW) from the idle state, based on either the MI-SW model or the MI-Ton model, may serve as a possible switch to trigger physical swallow which could be facilitated by traditional physical means with synchronization.

The objectives of this study were to: (1) test the hypotheses that MI-SW and MI-Ton could be detected from the background idle state for their possible use in stroke dysphagia rehabilitation; (2) build a model based on another simple yet relevant modality of MI-Ton electroencephalography (EEG) signals and use the model to detect MI-SW, (3) test the hypotheses by determining the classification accuracies across sessions and modalities, (4) determine the classification accuracies in a sample of ten healthy volunteers and one subject with chronic stroke. Some preliminary results of this work were presented in [13–15]. It should be noted that this paper focused on detecting MI-SW EEG signals for stroke dysphagia rehabilitation based on an MI-SW model or a simple yet relevant MI-Ton model. To achieve satisfactory classification accuracies across sessions and modalities, the model was adaptively selected such that the ratio of the between-classes distances versus within-class distances based on the features of training and evaluation data can be maximized, while our earlier work focused on proposing a model adaptation technique to address the non-stationarity problem [14], which was tested on MI-SW data. The model that yielded the largest number of consistent features between training and evaluation data was selected; however, separation between the two classes was not considered.

2. Materials and methods

2.1. Subjects

Two phases of study were carried out in the following manner: phase 1, a prospective study of a convenient sample of ten healthy subjects was conducted in a research lab; this was followed by phase 2, a single-case study of a chronic stroke dysphagia patient who participated in similar EEG-based experiments in an ambulatory rehabilitation clinic. Prior to recruitment, ethical approval was obtained from the hospital's institutional review board and written informed consents were obtained from all subjects. In phase 1 of the study, a total of ten healthy subjects (i.e., eight males and two females) with ages of 35.9 $\pm$ 7.7 years (mean $\pm$ standard deviation) participated in the experiments. None of the healthy subjects had a history of respiratory, swallowing or neurological disorders, or eating or nutrition problems. Seven subjects had experience of motor imagery of hand grasping and three subjects were brain–computer interface-naive. In phase 2 of the study, a stroke subject who was a 56 year old ethnic Chinese male participated in the experiment. He presented with severe brainstem hemorrhagic stroke involving the right hemipons and midbrain with mass effect on the fourth ventricle one year prior to enrollment. The etiology of the stroke was hypertensive in nature. He had tetraparesis and severe post-stroke dysphagia with complete dependence on nasogastric tube feeding. Two months following his stroke, fiberoptic endoscopic evaluation of swallowing showed the presence of moderate oropharyngeal dysphagia with reduced orolingual control. Delayed swallows, mild reduced hyo-laryngeal excursion and pharyngeal stripping were present. Increased aspiration risk was seen with more delayed swallows when the patient was fatigued, leading to an increased aspiration risk. He also had moderately severe cerebellar dysarthria.

2.2. The experimental design

Three sessions of experiments were conducted for healthy subjects with each session consisting of two runs. The first two sessions consisted of MI-SW and idle, whereas the third session consisted of MI-Ton and idle. The purposes of collecting two sessions of MI-SW data were to evaluate the performance of session-to-session classification of MI-SW, whereas only one session of MI-Ton data was collected to test classification of MI-SW using a simple yet relevant MI-Ton model. Session-to-session classification performance of MI-Ton was not the main concern of this paper. Two sessions of experiments were conducted for the stroke patient, with each session consisting of two runs: one run of MI-SW versus idle, and one run of MI-Ton versus idle. The detailed settings of

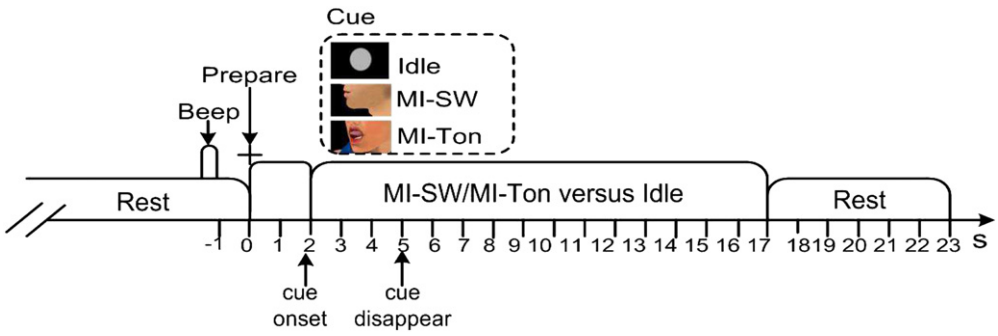


Figure 1. Timing scheme for EEG data collection.

Table 1. Detailed experimental settings.

Subj.	Sess./run	Tasks	No. of trials
Healthy controls	1/1	MI-SW versus idle	40 + 40
	1/2	MI-SW versus idle	40 + 40
	2/1	MI-SW versus idle	40 + 40
	2/2	MI-SW versus idle	40 + 40
	3/1	MI-Ton versus idle	40 + 40
	3/2	MI-Ton versus idle	40 + 40
Patient	1/1	MI-SW versus idle	40 + 40
	1/2	MI-Ton versus idle	40 + 40
	2/1	MI-SW versus idle	40 + 40
	2/2	MI-Ton versus idle	40 + 40

Note: Subj.: subjects. Sess.: session. No.: number.

the experiments can be found in table 1. In performing MI-SW, the subjects were advised to imagine swallowing a cup of water, or juice, or food such as a noodle or a bolus, whereas for MI-Ton, the subjects were advised to imagine protruding the tongue as far as possible, and as many times as possible. The subjects were advised not to perform any actions and not to close their eyes for the idle state. The resting state shown as a blue progress bar firstly appeared on the black screen. A short acoustic tone was then presented for 1 s, followed by a preparation of 2 s shown as a '+' symbol. Thereafter, the cue in the form of a virtual character performing swallowing for MI-SW, or tongue protrusion for MI-Ton, or a filled circle image for idle was shown for 3 s. With the disappearance of the visual cue, the subject started imagining the required tasks (for MI-SW or MI-Ton) or doing nothing (for idle) for 12 s. A rest time of 6 s was followed at the end of the trial. Each trial lasted for 23 s and the sequence of action and idle state in each run was randomized. The timing scheme of the experiments is shown in figure 1.

2.3. EEG and electromyography (EMG) recordings

The EEG measurements were obtained using the Neuroscan NuAmps EEG acquisition hardware [16], digitally subsampled at 250 Hz with a resolution of 22 bits and voltage ranges of ±130 mV. The placements of all 40 electrodes in the EEG cap followed the international 10–20 electrode placement standard. A notch filter of 50 Hz was enabled to remove artifacts caused by electrical power lines. In the experiments, the subjects were instructed to avoid physical movements such as eye blinking and unintentional swallow. To monitor

Figure 2. Placements of the electrodes for EMG recording.

the EMG activity during imagination or idle, EMG recording was done using two pairs of electrodes from the NuAmps taped beneath the skin of the submental and infrahyoid muscle groups, as shown in figure 2. The choice of these locations to measure the EMG activity was based on the facts that submental muscles are important in hyoid laryngeal excursion during swallowing. The electrodes were connected to the NuAmps EEG amplifier. Four channels, namely HEOL, HEOR, VEOU and VEOL, were used for EMG recordings. The other thirty-four channels, i.e., Fp1, Fp2, F7, F3, Fz, F4, F8, FT7, FT9, FC3, FCz, FC4, FT8, FT10, T7, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, O1, Oz, O2, PO1 and PO2, were used for EEG recordings. The average of the EEG signals for channels A1 and A2 (i.e., $\frac{A1+A2}{2}$) was used as the reference.

2.4. Dual-tree complex wavelet transform-based feature extraction (DTCWT-FE)

Our proposed feature extraction was based on the dual-tree complex wavelet transform (DTCWT) [17–19]. Sets of double-density dual-tree complex wavelets were employed [18, 19], which employed two scaling functions and four distinct wavelets. The wavelets were designed such that the two wavelets of the first pair were offset from one another by one half, and the other pair of wavelets formed an approximate Hilbert transform pair. In the implementation, the filter banks were constructed using finite impulse response perfect reconstruction filter banks, which were applied recursively to the low-pass sub-band using the analysis and synthesis filters for the forward and inverse transform [18]. The reasons for selecting this wavelet were due to its perfect reconstruction, being nearly shift-invariant, with approximately analytic real

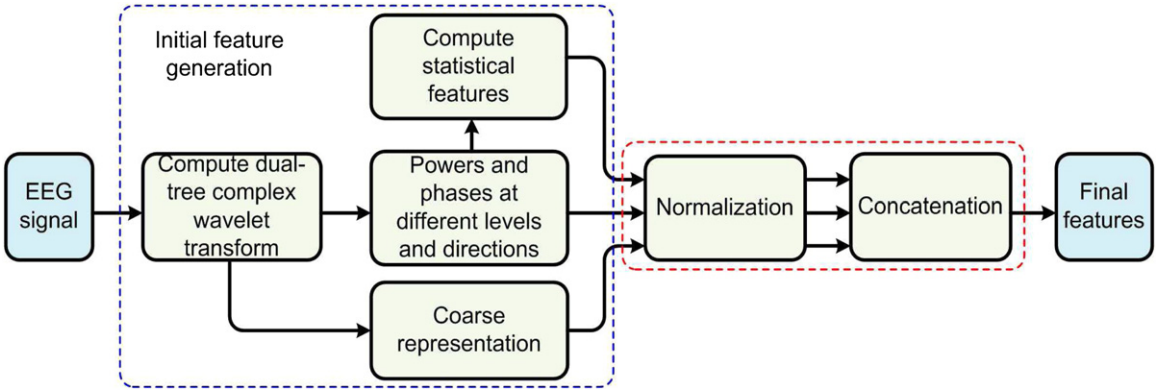


Figure 3. Schematic diagram illustrating the proposed DTCWT-FE scheme.

and imaginary parts of the filters and anti-aliasing effects [18, 19]. Further, the impulse responses of the filters were similar to that of the ERD/ERS produced during motor imagery. The 1D DTCWT decomposed and represented the input signal S as a complex shifted and dilated mother wavelet $\psi(x)$ and scaling function $\phi(x)$, which was given by

$$S(x) = \sum_{J\in Z, d\in\{1,2\}} C_{J+1,d}\phi_{J+1,d}(x) + \sum_{s\leqslant J}\sum_{d\in\{1,2\}} C_{s,d}\psi_{s,d}(x) \tag{1}$$

where Z, J, s and d denote the set of natural numbers, total decomposition level, level index and direction index, respectively. $C_{s,d}$ is the complex wavelet coefficient with $\psi_{s,d}(x) = \psi_{s,d,1}(x) + \sqrt{-1}\psi_{s,d,2}(x)$ and $\phi_{J+1,d}(x) = \phi_{J+1,d,1}(x) + \sqrt{-1}\phi_{J+1,d,2}(x)$. The signal was decomposed into $J = 5$ levels in the experiments. The highest frequency component in the signal was 125 Hz since the sampling frequency was 250 Hz. In this case, the *delta* rhythm was represented by coarse coefficients $C_{J+1,d,r}$, where r ($r \in \{1, 2\}$) denotes real and imaginary coefficients, while the *theta*, *alpha*, *beta* and *gamma* rhythms were represented by detailed coefficients $C_{s,d,r}$ at levels 5, 4, 3, and levels 2 and 1, respectively. To detect event-related synchronization/desynchronization (ERS/ERD) during motor imagery, the features consisted of the following at different levels and directions: powers (F_w), phases (F_p), coarse representation of the EEG signals (F_a), and higher order statistics such as skewness and kurtosis of the powers and phases. A schematic diagram illustrating the proposed dual-tree complex wavelet transform-based feature extraction scheme (DTCWT-FE) is shown in figure 3.

The power and phase at level s and direction d ($F^w_{s,d}$ and $F^p_{s,d}$), and coarse representation ($F^a_{s,d}$) are given by

$$F^w_{s,d} = \sum_{i=1}^{n_s} A^f_{s,d}(i) \tag{2}$$

$$F^p_{s,d} = \sum_{i=1}^{n_s} P^f_{s,d}(i) \tag{3}$$

$$F^a_{s,d} = C_{J+1,d,1} \tag{4}$$

where n_s denote the length of coefficients at level s and direction d; $A^f_{s,d}$ and $P^f_{s,d}$ were calculated by using

$$A^f_{s,d} = (C_{s,d,1})^2 + (C_{s,d,2})^2 \tag{5}$$

$$P^f_{s,d} = \arctan\left(\frac{C_{s,d,2}}{C_{s,d,1} + \epsilon}\right) \tag{6}$$

where ϵ is a small constant used to avoid the denominator being zero; $\arctan(x)$ is the arctangent of element x. The shape and tail direction information for the distribution of the coefficients is important in the detection of ERD/ERS generated during motor imagery; the 'skewness' and 'kurtosis' were computed to measure the 'asymmetry' and 'peakedness' of the distribution of powers and phases of the wavelet coefficients at different levels (s) and directions (d), which were given by

$$K^{kw}_{s,d} = S_{kw}(K^f_{s,d}) \tag{7}$$

$$K^{ur}_{s,d} = C_{ur}(K^f_{s,d}) \tag{8}$$

where the $K = \{A, P\}$ represent the amplitude and phase, respectively. The skewness and kurtosis were calculated by using

$$S_{kw}(x) = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}} \tag{9}$$

$$C_{ur}(x) = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{2}} - 3 \tag{10}$$

where $x = \{x_1, x_2, \ldots, x_i, \ldots\}$ and $\bar{x}$ is the sample mean. The nominator and denominator were the third central moments and sample variance, and the fourth and second sample moments about the mean, for equations (9) and (10), respectively. $F_w, F_p, F_a, A_{kw}, A_{ur}, P_{kw}$ and P_{ur} were obtained by concatenating $F^w_{s,d}, F^p_{s,d}, F^a_{s,d}, A^{kw}_{s,d}, A^{ur}_{s,d}, P^{kw}_{s,d}$ and $P^{ur}_{s,d}$ at different levels and directions, respectively. These features were further normalized by deducting the mean and divided by the standard deviation to obtain $\tilde{F}_w$, $\tilde{F}_p$, $\tilde{F}_a$, $\tilde{A}_{kw}$, $\tilde{P}_{kw}$, $\tilde{A}_{ur}$ and $\tilde{P}_{ur}$. Finally, the wavelet-based feature vector (F_v) was given by

$$F_v = \tilde{F}_w||\tilde{F}_p||\tilde{F}_a||\tilde{A}_{kw}||\tilde{P}_{kw}||\tilde{A}_{ur}||\tilde{P}_{ur} \tag{11}$$

where '||' denotes 'concatenation'. The features generated using the proposed DTCWT-FE for both MI-SW and MI-Ton are illustrated in figure 4. In the experiments, the EEG signals were decomposed into $J = 5$ levels. To illustrate the wavelet features, the electrodes 'C3' and 'C4' were selected, and an entire time segment from 2.5 s to 15.5 s after the onset of the visual cue was employed to compute the DTCWT.

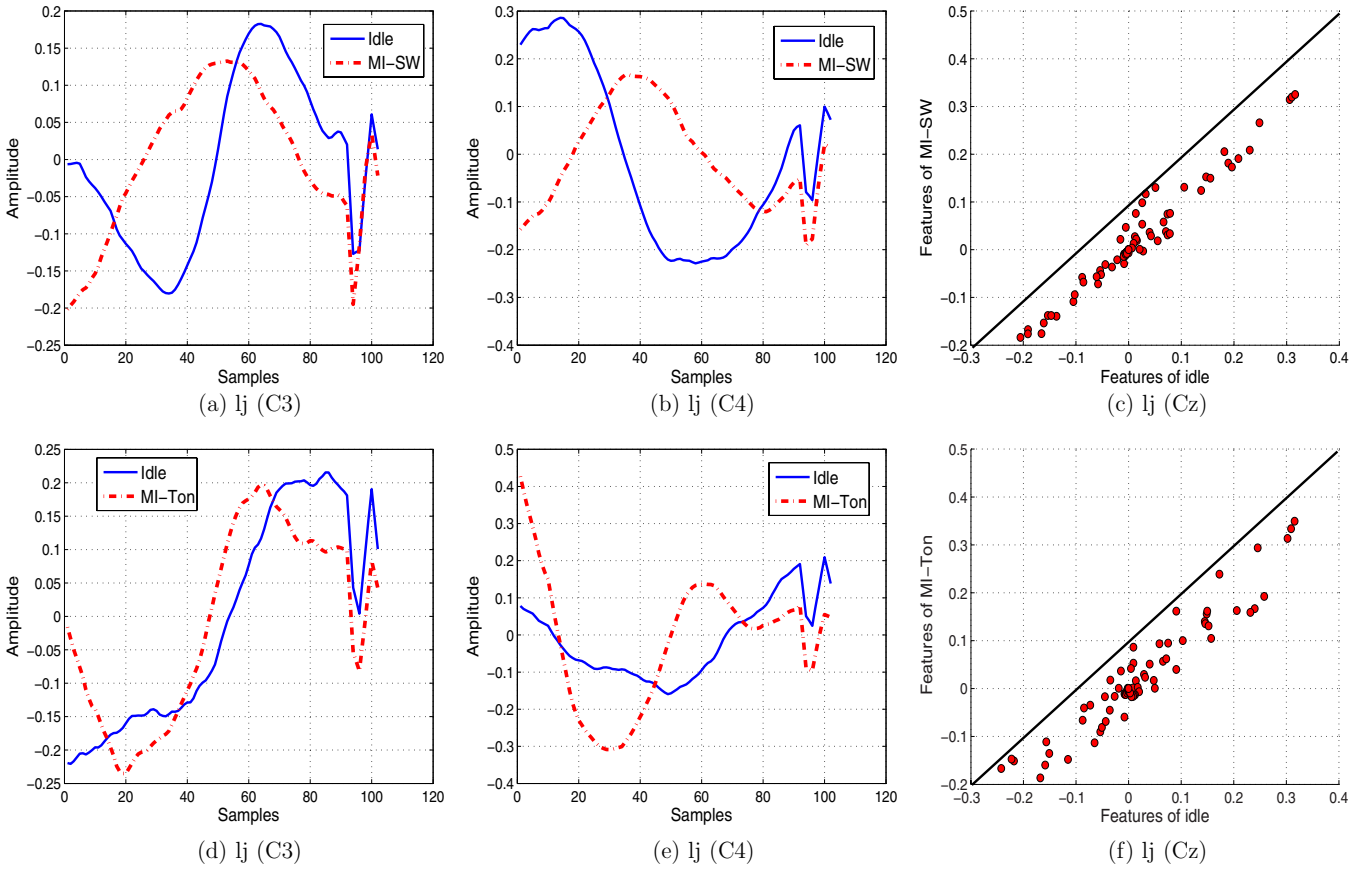


Figure 4. Illustration of the features generated by proposed DTCWT-FE method. (a), (b) and (d), (e): coarse representation of MI-SW versus idle, and MI-Ton versus idle EEG signals for subject 'lj'. (c) and (f): scatter plot of the whole set of features averaged across trials for MI-SW versus idle (c), and MI-Ton versus idle (f).

The features obtained for each class were averaged across all the trials in the same class, e.g., 80 trials of motor imagery and 80 trials of idle were used in the analysis. The coarse signal obtained was further smoothed with a moving window (e.g., of length 10) with the resultant coarse signal shown in figures 4(a), (b) and (d), (e). Significant ERD and ERS could be seen at the beginning and in the middle of the motor imagery from the coarse signal [20], while electrode 'Cz' was selected and the subject-specific optimal time segment of 2 s was employed to compute the entire set of features, with the scatter plot of features shown in figures 4(c) and (f).

2.5. Adaptive model selection for session-to-session classification

The assumption that training and testing data followed the same distribution in many machine learning algorithms [21] may not hold due to the non-stationarity of EEG signals. The non-stationarity was caused by the changes in electrode locations, variations in the electrodes impedance due to dry-up of electrodes over sessions, changes in mental states (e.g., fatigue and vigilance), different attention levels, and background noise and artifacts caused by eye and muscle movements. The visual feedbacks during online feedback sessions could also contribute to the non-stationarity. Existing methods for tackling the non-stationarity include covariate shift adaptation [22, 23]. The effects of non-class-related non-stationarities in motor imagery EEG were analyzed; subsequently, the parameters of a linear classifier without label information were adapted [23]. 'ReTrain' and 'ReBias' were proposed for classifier adaptation to the testing data [24]. The continuous output of the modified classifier was used in 'ReBias', which was shifted by a certain amount to minimize the error on the labeled feedback data, whereas the offline features were employed in 'ReTrain' to train the linear discriminant analysis classifier; subsequently, the hyperplane that yielded the minimum error on labeled feedback data was chosen.

Our idea was to select the most suitable model from the multiple training models generated during cross-validation, for the classification of the given testing data. Specifically, the model was selected by maximizing ratio of the between-classes distances and within-class distances (**msBWD**), evaluated via the features of the training and evaluation data. Hence, the testing data consisted of two parts: a small amount of evaluation data used for model selection [25], and the testing data used to test the session-to-session classification performance.

2.5.1. Calculation of within-class and between-classes distances. The suitability of each model was evaluated on the basis of the proposed within-class and between-classes distances calculated using the features of training and

evaluation data. Let us denote the classifier as C_r and assume that an r times and n-fold cross-validation was carried out; the $(p, q)th$ generated model was given by

$$M_d(p, q) = C_r(F_v, I_{tr}(p, q), I_{te}) \tag{12}$$

where $p = 1, 2, \ldots, r$ and $r = 1, 2, \ldots, n$; I_{tr} and I_{te} denote the indexes of the training data in cross-validation and evaluation data, respectively. The models obtained on the basis of these random partitions were used for subsequent model selection catering for the testing data. For ease of presentation, we assumed the balance of trials in the two classes. Let us denote the feature vectors for training and evaluation data as F_{tr} and F_{te}, respectively, the number of trials for training and evaluation data as $2m$ and $2n$, respectively, and the dimensionality of the feature vector as k. The features for class 'i' of training (F_{tr}^i) and evaluation data (F_{te}^i) were given by

$$F_{tr}^i = F_{tr}|(Y_{tr}(I_{tr}) = i) \tag{13}$$

$$F_{te}^i = F_{te}|(Y_{te}(I_{te}) = i) \tag{14}$$

where I_{tr} and Y_{tr}, and I_{te} and Y_{te} are the indexes and labels for training and evaluation data, respectively, and $i \in \{0, 1\}$ represents classes '0' and '1'. The *normalized correlation* distances of within-class (D_{00} and D_{11}) and between-classes (D_{10} and D_{01}) cases were calculated by using

$$D_{ij} = D_s(F_{tr}^i, F_{te}^j) \tag{15}$$

where $i, j \in \{0, 1\}$ represent classes '0' and '1'. $D_s(\mathbf{a}, \mathbf{b})$ was used to calculate the normalized correlation distances between feature vectors $\mathbf{a}$ and $\mathbf{b}$, given by

$$D_s(\mathbf{a}, \mathbf{b}) = \frac{A_b}{\sqrt{A_a} * \sqrt{B_b}} \tag{16}$$

where A_a, B_b and A_b are given by $A_a = \sum_{j=1}^{k} (a_{uj}. * a_{uj})$, $B_b = \sum_{j=1}^{k} (b_{vj}. * b_{vj})$ and $A_b = \sum_{j=1}^{k} (a_{uj} * b_{vj}^{\mathrm{T}})$, where X^{T} represents the matrix transpose of X; A_a, B_b and A_b were of dimension $u \times 1$, $1 \times v$ and $u \times v$, respectively. Obviously, the size of D_{ij} was $m \times n$, considering the balance of the two classes. Removal of the respective mean distances from the within-class distances and between-classes distances was then carried out to obtain D_{ii}^w and D_{ij}^b $(i \neq j)$, i.e., $D_{ii}^w = D_{ii} - \bar{D}_{ii}$ and $D_{ij}^b = D_{ij} - \bar{D}_{ij}$ where $\bar{D}_{ii}$ and $\bar{D}_{ij}$ are the mean distances of within-class $(i = \{0, 1\})$ and between-classes $(i \neq j,$ $i, j \in \{0, 1\})$ cases.

2.5.2. The criterion for model selection. The covariance matrices of the within-class distances and between-classes distances were represented as C_{ii} and C_{ij} $(i \neq j)$, respectively, which were calculated on the basis of D_{ii}^w and D_{ij}^b and given by

$$C_{ii} = \frac{1}{n}(D_{ii}^w)^{\mathrm{T}} D_{ii}^w \tag{17}$$

$$C_{ij} = \frac{1}{n}(D_{ij}^b)^{\mathrm{T}} D_{ij}^b \tag{18}$$

Thereafter, the distances to the hyperplane w were calculated by using

$$w = \mu_d * \Sigma_w^{-1} \tag{19}$$

where μ_d and Σ_w are given by

$$\Sigma_w = \frac{1}{2}\left(\sum_{i=\{0,1\}} C_{ii} + \sum_{\{i,j\}\in\{0,1\}, i\neq j} C_{ij}\right) \tag{20}$$

$$\mu_d = \frac{1}{2}\left(\sum_{i=\{0,1\}} \bar{D}_{ii} - \sum_{\{i,j\}\in\{0,1\}, i\neq j} \bar{D}_{ij}\right). \tag{21}$$

Subsequent projections of the within-class and between-classes distances (σ_w and σ_b) into the direction of w were given by

$$S_w = w^{\mathrm{T}} * \sigma_w * w \tag{22}$$

$$S_b = w^{\mathrm{T}} * \sigma_b * w \tag{23}$$

where σ_b and σ_w are given by

$$\sigma_b = \frac{1}{2}\left(\sum_{\{i,j\}\in\{0,1\}, i\neq j} D_{ij}^b D_{ij}^{b\,\mathrm{T}}\right) \tag{24}$$

$$\sigma_w = \frac{1}{2}\left(\sum_{i=\{0,1\}} D_{ii}^w D_{ii}^{w\mathrm{T}}\right). \tag{25}$$

Finally, the best separation of the within-class and between-classes distances in the direction of w was given by

$$(V_e, U_e) = \mathrm{eigs}\left(\frac{S_b}{S_w}\right) \tag{26}$$

$$R_{cd}(p, q) = trace(U_e) \tag{27}$$

where eigs() was the method used for solving the generalized eigenvalue problem, e.g., $AV_e = U_e A$ for matrix A and $A = \frac{S_b}{S_w}$. U_e was the diagonal matrix of $\frac{S_b}{S_w}$'s eigenvalues and for matrix V_e the columns were the corresponding eigenvectors; trace(X) was the trace of matrix X, e.g., trace$(X) = \sum_{i=1}^{k} U_e(i)$, which was the sum of the eigenvalues. The largest eigenvalue corresponded to the maximum separation of $\frac{S_b}{S_w}$. Hence, the $(\hat{p}, \hat{q})$th model was selected by using

$$(\hat{p}, \hat{q}) = \underset{p,q}{\mathrm{argmax}}(R_{cd}(p, q)) \tag{28}$$

A schematic diagram illustrating the session-to-session classification based on our proposed model selection criterion is shown in figure 5. Once the model was selected on the basis of our proposed criterion, i.e., **msBWD**, the features of the testing data from a moving window of 2 s were classified. Finally, the best accuracy among all the moving windows was reported as the testing accuracy.

3. Results

A total of 160 trials for healthy controls (i.e., 80 trials of motor imagery and 80 trials of idle), and 80 trials for a stroke patient (i.e., 40 trials of motor imagery and 40 trials of idle) in a session were used in the performance evaluation of cross-validation classification. Further, the features of 160 trials for healthy controls and 80 trials for a stroke patient from one session of a subject were used to train the model, which was subsequently used for classifying the features of the same number of trials from another session for the same subject, in the performance evaluation of session-to-session classification. No trial was excluded in the experiments; details of the trials can be found in table 1.

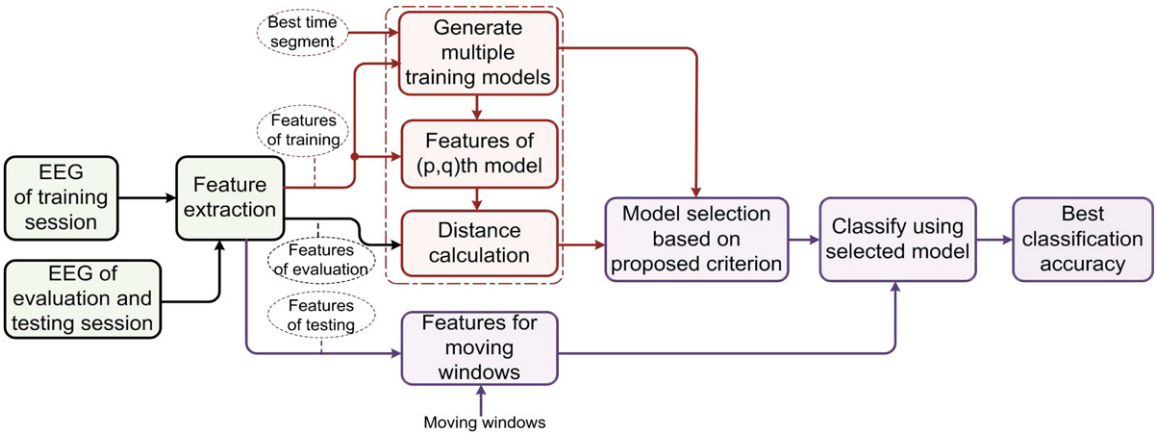


Figure 5. Schematic diagram illustrating the session-to-session classification using the proposed model selection based on between-classes and within-class distances (msBWD).

3.1. Cross-validation classification accuracies

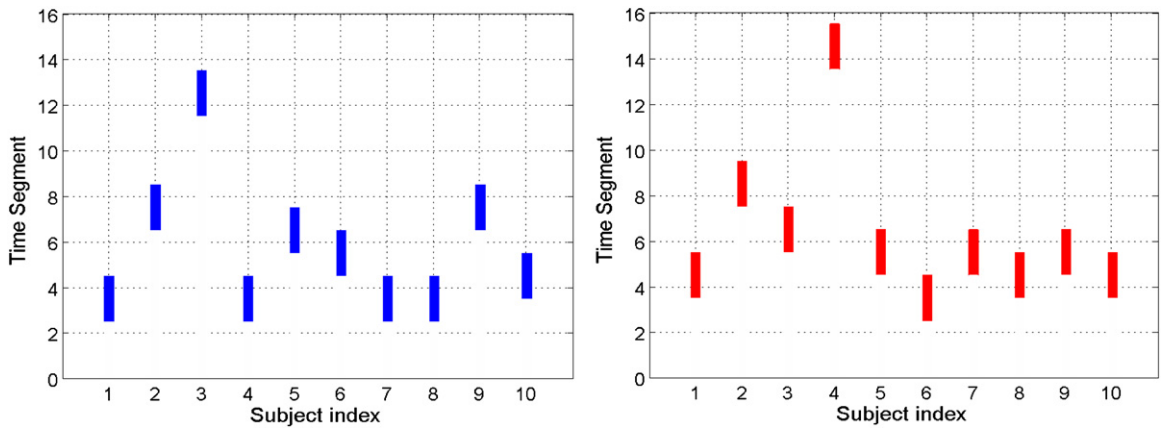
We firstly evaluate the classification performance on the basis of 10 × 10-fold CV. The indexes for a 10-fold CV for all of the observations were randomly generated each time. In each part, the current part was reserved for validation, while the other nine parts from random partition were used to train the model. This would repeat for ten times with the mean accuracy being reported. Two frontal EEG channels of FP1 and FP2 related to EOG artifacts and eye blinks, and two side EEG channels of FT9 and FT10 were excluded. The support vector machine (SVM) is a powerful classification tool that classifies data by finding the best hyperplane to separate all data points of one class from those of the other classes. The SVM has demonstrated its superior performance for the classification of EEG signals and has been the state-of-the-art model for other classification tasks [9, 13–15, 26–28, 32]. Since the feature dimension in our method was much larger than the number of data instances, the 'rbf' kernel was not suitable [28]; hence, a 'linear' kernel SVM was selected as the classifier. In the implementation, the bioinformatics toolbox of Matlab was employed and the regularization factor (denoted as C) was set to the default value of '1'.

3.1.1. Comparison of cross-validation accuracies using different time segments. The accuracies of classifying MI-SW and MI-Ton from the background idle state using our proposed DTCWT-FE scheme were compared with those obtained using other methods such as the: common spatial pattern (CSP) [29, 30]; filter bank CSP (FBCSP) [31]; and sliding window discriminative CSP (SWDCSP) [32]. The subject-specific optimal time segments of 2 s were manually selected for our proposed method by studying the motor imagery of the swallow pattern. A small amount of trials can be employed to derive the best time segments on the basis of a nested cross-validation in a supervised manner. This is applicable for MI-Ton since it is simple and the activation patterns are relatively stable. In the experiments, 20 trials of MI-Ton and 20 trials of idle were employed to select the best time segment that yielded the highest accuracy in cross-validation. However, a subject-specific stable time segment for MI-SW was difficult to obtain without using sufficient trials, e.g., 80 trials. This was due to the long time required to finish the motor imagery of swallow. Large variation existed in the time to reach ERD/ERS, as shown in figure 6 for the time segments selected using CV. The time segments from −0.5 s to 1.5 s after the onset of the action cue were selected for other methods, with the comparison results for MI-SW and MI-Ton shown in tables 2 and 3. As observed from the tables, our proposed method achieved the best accuracies among all the methods. The averaged accuracies across subjects for the classification of MI-SW

Table 2. Comparison of cross-validation classification accuracies of MI-SW versus idle using proposed and other methods.

sb./	Cross-validation accuracy ($\mathcal{A}_c \pm \mathcal{V}_r$)			
ss.	DTCWT-FE[a]	FBCSP[b]	CSP[c]	SWDCSP[d]
sd/1	**61.10** ± 3.64	52.31 ± 2.96	51.38 ± 2.43	46.05 ± 3.36
sd/2	**63.19** ± 3.35	55.94 ± 3.80	60.63 ± 2.08	54.94 ± 2.52
lj/1	71.25 ± 2.49	**76.94** ± 1.94	69.25 ± 3.06	76.24 ± 1.84
lj/2	**82.37** ± 3.54	61.81 ± 2.53	67.31 ± 1.35	62.40 ± 1.81
hj/1	80.59 ± 1.96	**81.31** ± 2.09	78.75 ± 1.41	71.00 ± 2.92
hj/2	84.07 ± 2.42	**88.31** ± 1.56	84.13 ± 2.42	86.61 ± 2.26
aw/1	**64.19** ± 3.70	54.81 ± 3.63	62.75 ± 3.14	52.92 ± 4.51
aw/2	**62.44** ± 3.34	57.00 ± 2.33	61.44 ± 2.44	55.88 ± 4.40
cr/1	**84.69** ± 2.17	73.06 ± 3.14	81.88 ± 2.10	62.83 ± 3.01
cr/2	**96.67** ± 1.18	90.19 ± 1.59	91.94 ± 4.14	83.81 ± 3.23
wy/1	76.61 ± 2.36	**80.75** ± 1.76	76.00 ± 1.57	78.88 ± 3.47
wy/2	**76.90** ± 0.85	74.75 ± 1.82	65.44 ± 2.82	72.72 ± 3.58
cc/1	**63.67** ± 4.69	51.00 ± 3.80	53.63 ± 2.23	52.75 ± 2.55
cc/2	69.36 ± 2.94	**77.38** ± 2.12	75.06 ± 2.17	74.02 ± 1.71
mt/1	**68.00** ± 2.78	64.69 ± 2.11	57.50 ± 3.02	62.48 ± 4.66
mt/2	60.43 ± 5.23	**62.56** ± 1.54	57.94 ± 2.70	59.01 ± 3.50
zy/1	**67.12** ± 2.95	60.31 ± 2.95	64.81 ± 2.00	60.64 ± 3.59
zy/2	**59.59** ± 2.59	50.69 ± 3.45	58.75 ± 3.92	53.98 ± 1.13
cj/1	**60.01** ± 3.37	51.25 ± 3.19	51.56 ± 3.70	49.97 ± 4.07
cj/2	**65.59** ± 2.60	51.75 ± 2.24	50.94 ± 2.81	53.22 ± 2.60
A_{as}	**70.89** ± 2.91	65.84 ± 2.53	66.05 ± 2.58	63.52 ± 3.04
Pt.		a versus b	a versus c	a versus d
Pv.		**0.0066	**0.000 89	**0.000 31

A_c: accuracy (%), $\mathcal{V}_r$: variance; A_{as}: averaged accuracy across subjects and sessions; sb./ss.: subject/session; Pv.: p-value; Pt.: paired t-test. The best performance is shown in bold.
**: significant.



(a) session 1 (b) session 2

Figure 6. Best time segments selected by cross-validation for motor imagery of swallow versus idle for two sessions.

Table 3. Comparison of cross-validation classification accuracies of MI-Ton versus idle using proposed and other methods.

sb./	Cross-validation accuracy ($\mathcal{A}_c \pm \mathcal{V}_r$)			
ss.	DTCWT-FE[a]	FBCSP[b]	CSP[c]	SWDCSP[d]
sd/1	**62.43** ± 3.09	48.75 ± 3.39	59.81 ± 2.78	50.93 ± 1.96
lj/1	84.77 ± 2.11	**85.56** ± 2.01	82.56 ± 2.39	84.47 ± 1.53
hj/1	81.10 ± 3.07	79.44 ± 3.34	**81.50** ± 2.23	70.61 ± 3.49
aw/1	**63.77** ± 3.22	47.06 ± 3.64	52.69 ± 2.42	48.02 ± 2.01
cr/1	97.16 ± 0.99	97.00 ± 0.26	**98.50** ± 1.40	93.05 ± 1.84
wy/1	**83.48** ± 2.02	81.13 ± 2.06	78.69 ± 2.67	76.99 ± 3.50
cc/1	63.08 ± 1.62	65.38 ± 2.57	62.25 ± 3.45	**67.78**± 1.45
mt/1	**75.60** ± 1.77	63.00 ± 1.17	65.00 ± 2.31	63.56 ± 2.72
zy/1	**63.43** ± 2.82	57.44 ± 3.18	56.25 ± 2.64	56.44 ± 3.37
cj/1	**63.12** ± 4.57	56.19 ± 3.91	50.94 ± 2.51	50.18 ± 2.25
A_{as}	**73.79** ± 2.53	68.10 ± 2.55	68.82 ± 2.48	66.20 ± 2.41
Pt.		a versus b	a versus c	a versus d
Pv.		**0.0241	**0.0118	**0.0041

Table 4. Cross-validation classification accuracies of MI-SW versus idle and MI-Ton versus idle for the stroke dysphagia patient.

Tasks	sb./ss.	Cross-validation accuracy ($\mathcal{A}_c \pm \mathcal{V}_r$)
MI-Ton versus idle	sp/1	78.28 ± 2.39
	sp/2	62.19 ± 5.19
	A_s	70.24 ± 3.79
MI-SW versus idle	sp/1	54.93 ± 5.37
	sp/2	77.86 ± 4.47
	A_s	66.40 ± 4.92

A_s: averaged accuracy across sessions (%).

versus idle were significantly higher than those obtained using FBCSP (5.05%, p-value: 6.6×10^{-3}), CSP (4.84%, p-value: 8.9×10^{-4}) and SWDCSP (7.37%, p-value: 3.1×10^{-4}) at the 5% significance level. Similarly, the averaged classification accuracies across subjects for the classification of MI-Ton versus idle were significantly higher than those obtained for FBCSP (5.69%, p-value: 0.0241), CSP (4.97%, p-value: 0.0118) and SWDCSP (7.59%, p-value: 0.0041). The higher averaged CV accuracies for MI-Ton in comparison to those for MI-SW revealed that it was easier to perform MI-Ton than MI-SW for stroke rehabilitation. The CV accuracies for the stroke dysphagia patient (denoted as: 'sp') are shown in table 4. The averaged accuracies of 66.40% and 70.24% were obtained for MI-SW versus idle and MI-Ton versus idle. Like for the results obtained for healthy subjects, the CV accuracies for MI-Ton were higher than those for MI-SW.

Strong correlation was observed between the classification accuracies of MI-SW versus idle and MI-Ton versus idle across subjects. The Pearson correlation coefficients were 0.9143 (p-value: 2.1263×10^{-4}) and 0.8843 (p-value: 6.7991×10^{-4}) between the accuracies of MI-SW of session 1 and MI-Ton, and between the accuracies of MI-SW of session 2 and MI-Ton, respectively. The strong correlation not only revealed the subjectwise similar responses to the two motor imagery tasks, but also showed the possibility of classifying MI-SW using the MI-Ton model, with the result that a switch could be provided to trigger actual swallow for rehabilitation. Estimation of the accuracy at 95% confidence on the respective action performed at chance level was done using the binomial inverse cumulative distribution function for a total of 80 and 160 trials (for the stroke patient and healthy subjects) in one session. The results indicated that the subjects whose accuracy fell between 38.75% and 61.25% (for the stroke patient) and between 42.5% and 57.5% (for healthy subjects) were deemed as performing at chance level. Hence, all healthy subjects performed above chance level using the proposed method. However, the stroke patient performed at the chance level for MI-SW session 1.

3.1.2. Comparison of cross-validation accuracies using the same subject-specific time segments. To exclude the confounding effects caused by the difference in selecting time segments, we also carried out the comparison of accuracies of the proposed and other methods on the basis of the same selected subject-specific time segments. The pairwise scatter plot of the cross-validation accuracies for different pairs of methods is shown in figure 7. By employing the selected subject-specific time segments, the averaged accuracies (%) of MI-SW obtained for DTCWT-FE, CSP, FBCSP and SWDCSP were: 70.89 ± 2.91, 67.95 ± 2.34, 68.02 ± 2.10 and 66.58 ± 2.71, respectively. The averaged accuracies of CSP,

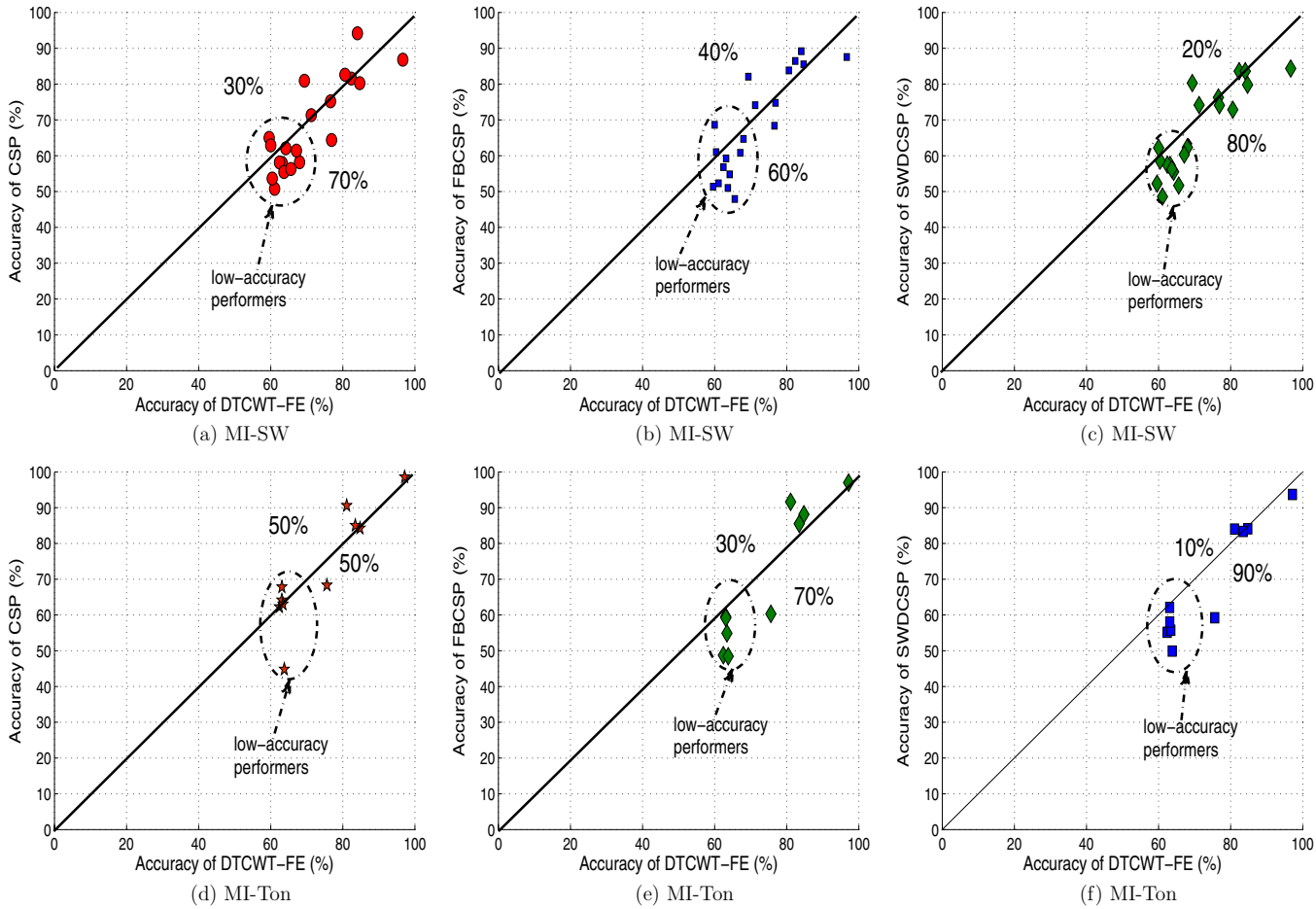


Figure 7. Pairwise scatter plot of CV classification accuracies using the same selected time segments for (a)–(c): MI-SW and (d)–(f): MI-Ton. ((a), (d)) DTCWT-FE and CSP; ((b), (e)) DTCWT-FE and FBCSP; and ((c), (f)) DTCWT-FE and SWDCSP. Each dot represents the classification accuracy of a subject.

FBCSP and SWDCSP were improved by 1.90%, 2.18% and 0.38%, compared with those obtained using the fixed time segments. However, the averaged accuracy of our proposed DTCWT-FE was still 2.94%, 2.87% and 4.31% higher than those obtained by CSP, FBCSP and SWDCSP, respectively. Similarly, compared with those obtained using the fixed time segments, the accuracies of MI-Ton using the selected time segments were improved by: 4.05%, 1.24% and 2.33% for CSP, FBCSP and SWDCSP, respectively. However, the averaged accuracies of our proposed method were still 0.92%, 4.45% and 5.26% higher than those for CSP, FBCSP and SWDCSP, respectively, using the selected time segments. Observing figure 7, it can be seen that our proposed method was especially effective for the low-accuracy performers, e.g., those subjects whose accuracies were below 70%. In contrast, the performance for the low-accuracy performers was poor for CSP-based methods due to the noise sensitivity of these methods, as supported by the results presented in tables 2 and 3, and figure 7.

3.2. Classification accuracies across sessions and modalities

The session-to-session classification accuracies of motor imagery of swallow EEG signals were evaluated using two kinds of models generated on the basis of the EEG signals from a different session: (a) MI-SW versus idle and (b) MI-Ton versus idle. The purpose was to test the session-to-session classification accuracies not only using the model generated from the same modality, but also using the model generated from a different yet similar modality for the purpose of rehabilitation of stroke patient. Note that the models were generated on the basis of the subject-specific time segments selected by cross-validation. Six healthy subjects (‘lj’, ‘hj’, ‘cr’, ‘wy’, ‘cc’ and ‘mt’) whose cross-validation classification accuracies for the three sessions (two sessions of MI-SW versus idle, and one session of MI-Ton versus idle) were above 63% were selected for the evaluation. 13 overlapping windows of 2 s moving through the whole time interval of [2.5, 15.5] s were employed; the best performance among the windows was reported as the test accuracy. In the experiments, the classification accuracies using the *full model* (i.e., training the model using the full set of features of training data), the *best model* (i.e., selecting the training model with the best cross-validation accuracy), and the selected model, based on our proposed *msBWD*, were compared with the results presented in table 5. The features of the evaluation data were further grand averaged across trials to improve robustness.

The results shown in table 5 revealed that the averaged session-to-session classification accuracies of 70.32% and 69.74% were achieved using the full MI-SW and MI-Ton

Table 5. Session-to-session classification accuracies of motor imagery of swallow EEG signals for healthy subjects.

sb./	Full model		Best model		msBWD (sw)			msBWD (ton)		
ss.	sw[a]	ton[b]	sw[c]	ton[d]	20tr.[e]	40tr.[f]	r-tn[g]	20tr.[h]	40tr.[i]	r-tn[j]
lj1	**75.00**	70.00	**75.00**	70.63	72.86	73.33	**75.00**	72.86	73.33	74.17
lj2	76.25	81.25	76.25	81.25	77.86	**78.33**	75.00	80.00	81.67	85.00
hj1	76.25	71.88	76.88	73.13	74.29	**77.50**	75.83	72.86	75.00	75.83
hj2	71.88	68.75	73.13	71.25	70.71	73.33	**76.67**	72.86	73.33	67.50
cr1	78.13	80.00	79.38	82.50	83.57	**86.67**	**86.67**	85.00	86.67	82.50
cr2	80.63	95.00	83.13	95.63	86.43	85.83	**92.50**	97.86	97.50	97.50
wy1	73.75	64.38	71.88	62.50	70.00	67.50	**77.50**	61.43	61.67	60.83
wy2	68.13	64.38	66.88	65.00	**68.57**	66.67	65.00	67.14	66.67	71.67
cc1	59.38	58.75	60.00	59.38	60.71	**61.67**	58.33	60.71	60.83	60.00
cc2	65.00	62.50	63.13	66.25	**67.14**	65.83	62.50	65.00	65.00	62.50
mt1	61.25	59.38	60.63	57.50	**65.71**	63.33	60.00	60.00	63.33	63.33
mt2	58.13	60.63	55.63	61.88	60.71	**62.50**	61.67	60.71	60.00	54.17
A_{as}	70.32	69.74	70.16	70.58	71.55	71.87	**72.22**	71.37	72.08	71.25
Pt.			a versus c	b versus d	a versus e	a versus f	a versus g	b versus h	b versus i	b versus j
Pv.			$0.73^{\times\times}$	$0.11^{\times\times}$	$0.19^{\times\times}$	$0.17^{\times\times}$	$0.19^{\times\times}$	0.03^{**}	0.01^{**}	$0.20^{\times\times}$

Note: r-tn: retrain; A_{as}: averaged accuracy across subjects and sessions (%); 20tr.: using 20 trials to select the model. The best performances using MI-SW (sw) and MI-Ton (ton) models are shown in bold and underlined, respectively. $^{\times\times}$: insignificant; $**$: significant.

Table 6. Session-to-session classification accuracies of motor imagery of swallow for the stroke patient.

	Full model		Best CV model		msBWD (sw)		msBWD (ton)	
Session	sw	ton	sw	ton	20tr.	40tr.	20tr.	40tr.
1	61.25	57.50	62.50	60.00	**65.00**	62.50	60.00	60.00
2	65.00	66.25	62.50	76.25	65.00	**67.50**	75.00	80.00
A_s	63.13	61.88	62.50	68.13	**65.00**	**65.00**	67.50	70.00

A_s: averaged classification accuracy across sessions.

models to classify MI-SW, respectively. The classification accuracies were further increased by 1.63% and 2.34% with the selection of the suitable MI-Ton model using 20 and 40 trials of evaluation data based on the proposed msBWD method. The increase was significant at 5% significance level with p-values of 0.029 and 0.007. However, retraining the model using the wavelet features from the selected model and those of the evaluation data did not further improve the accuracy, despite the fact that it was still 1.51% higher than that obtained using the full model. Similarly, the accuracy was increased by 1.23%, 1.56% and 1.91% by using 20 and 40 trials of evaluation data to select the MI-SW models for the classification, and retraining the model, respectively. However, this was not significant compared with the results of using the full model at 5% significance level. The reason for the performance drop on using the MI-Ton model to classify MI-SW for some subjects such as 'wy' lay in different strategies being implemented for MI-SW and MI-Ton. For example, during the imagination of MI-SW, no obvious tongue movements were involved; hence the session-to-session classification accuracy of using the MI-Ton model was not good. The paired t-test also revealed no significant difference on the classification accuracies for using MI-SW and MI-Ton models. The session-to-session classification accuracies of MI-SW using MI-SW and MI-Ton models for the stroke patient are shown in table 6. It can be seen that comparable performances were achieved using the full model of MI-Ton (accuracy: 61.88%) and using MI-SW (accuracy: 63.13%). The accuracy of using the MI-Ton model had been further boosted by 5.62% and 8.12% using 20 and 40 trials to select the model, respectively. However, on selecting a suitable MI-SW model for classifying MI-SW, the accuracies only increased by 1.87% and 2.5% compared with those achieved using the full model and best model, respectively.

4. Discussion and conclusions

To summarize, in this paper, two hypotheses were tested and the results revealed that: (1) MI-SW and MI-Ton could be detected from background idle for ten healthy subjects and one stroke dysphagia patient; (2) MI-SW could be detected using a model built from MI-Ton, as evidenced by the similar session-to-session classification accuracies obtained using the MI-SW model and MI-Ton model. The proposed dual-tree complex wavelet transform-based feature extraction demonstrated its effectiveness in detection of MI-SW and MI-Ton, as supported by the averaged cross-validation classification accuracies, i.e., 70.89% and 73.79% for MI-SW and MI-Ton for healthy subjects, and 66.40% and 70.24% for MI-SW and MI-Ton for the stroke dysphagia patient. The averaged CV classification accuracies for both MI-SW and MI-Ton for healthy subjects were significantly better than those obtained from CSP, SWDCSP and FBCSP. In particular, our proposed method performed especially well for the low-accuracy performers, thanks to the stable multi-level dual-tree complex wavelet features. In contrast, the

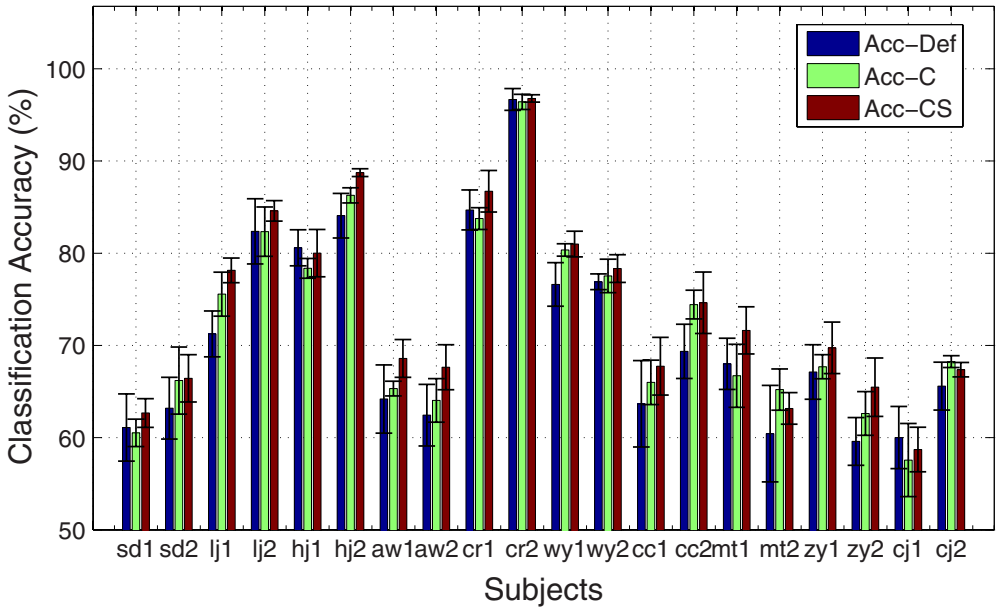


Figure 8. Comparisons of classification accuracies obtained by tuning C for 'linear' SVM ('Acc-C'), obtained by tuning both C and σ for 'rbf' SVM ('Acc-CS') and obtained by 'linear' SVM with the default regularization parameter ('Acc-Def').

performances of the CSP-based methods (e.g., CSP, FBCSP and SWDCSP) were easily affected by the noise, especially for low-accuracy performers. Statistical tests with 95% confidence estimating the accuracy on respective actions at the chance level showed that all subjects performed above chance level except for the stroke patient for MI-SW session 1, whose performance fell at the chance level. Further, the averaged session-to-session classification accuracies obtained for ten healthy subjects (one stroke patient) were 72.28% and 72.08% (65% and 70%) on selecting the best MI-SW model and MI-Ton model using the proposed model selection method.

While previous studies have documented comparable accuracies of motor imagery of hand/arm classification for healthy and stroke patients [29, 31, 32], our findings imply that obtaining similar classification accuracies is possible for the detection of MI-SW and MI-Ton for healthy subjects, despite the facts that performing MI-SW was more difficult than other modalities of motor imagery. This was due to the involvements of sensory processing, oral, pharyngeal and laryngeal movement execution, coordination of mastication and respiration and attentional processing [12], which complicated the process and led to high variability of the EEG signals among sessions and subjects. In contrast, other modalities of motor imagery including MI-Ton were relatively simple to perform, which motivated the use of the MI-Ton model for detecting MI-SW, as supported by the overlapping activated areas of swallowing and tongue movements [6, 12]. Comparing with healthy subjects, it was noticed that the averaged cross-validation accuracy of the single stroke dysphagia patient had dropped for both MI-SW (4.49%) and MI-Ton (3.55%); and 7.28% and 2.08% decreases in accuracies on selecting the best MI-SW model and MI-Ton model resulted. This agreed with previous findings of significant reductions of cortical sensorimotor activations in amyotrophic lateral sclerosis patients [33]. The lesser decrease in cross-validation accuracy for MI-Ton in comparison to MI-SW, and session-to-session classification accuracy of MI-SW using the MI-Ton model in comparison to the MI-SW model, further demonstrate the feasibility of using the MI-Ton model to detect MI-SW.

Note that 'linear' SVM was selected, with the regularization parameter being set to the default value in our method. Let us now discuss how the performance will be affected by selecting the parameters on the basis of 5×5-fold CV and grid search [28]. Selections of the parameters were carried out in two stages. Firstly, the regularization parameter (C) for 'linear' SVM, or the pair of parameters of (C, σ) for 'rbf' SVM that yielded the best performance were selected. In the implementation, C varying from 1×10^{-3} to 1×10^{6}, and σ varying from 1×10^{-5} to 1×10^{5}, increases by a factor of 10 were employed. A finer search was then implemented to select among the ten values centered at the parameters selected at the first stage. Comparisons of the classification performance by tuning the parameters with that obtained by 'linear' SVM with the default regularization parameter are shown in figure 8. The results revealed that the averaged CV accuracies were improved by 1.37% (p-value: 0.0148) and 3.01% (p-value: 0.0004) by tuning C for a 'linear' kernel, and by tuning both C and σ for the 'rbf' SVM, compared with those obtained for 'linear' SVM with the default C. However, it was noted that the selected best parameters vary significantly across sessions and subjects, which may create problems in implementation. Hence, we only reported the results from using 'linear' SVM with the default regularization parameter.

Non-cerebral artifacts produced by eye movements and eye blinks, and muscle movements, obscure and complicate the interpretation of EEG signals [34–36]. In the experiments, the subjects sat comfortably in a chair to perform the motor imagery tasks following the visual cue shown at the center of a computer screen. The vertical eye movements were

the primary artifacts reflected in FP1 and FP2, which were excluded in the processing to minimize the influence of confounding factors. FT9 and FT10 located at the most distal locations with respect to the center of the brain were easily affected by muscle artifacts and were not the primary activation regions for MI-SW and MI-Ton. Excluding these two channels for processing did not yield degradation of the classification performance. This was supported by the CV accuracies of using 32 EEG channels including FT9 and FT10, which were slightly lower than those achieved using 30 EEG channels without including these two channels. The EMG signals were monitored and manually inspected by an EEG expert to ensure that no excessive body movements, especially for unconscious swallowing and chewing, presented during the experiments. No trials were excluded in the processing.

The limitation of this exploratory study was that the sample size, i.e., ten healthy subjects and one dysphagia stroke patient, was relatively small. This limitation may serve as an impetus for a larger study to validate the important preliminary findings presented in this paper. The probable clinical efficacy for MI-SW or MI-Ton for swallowing rehabilitation could be targeted at common problems such as delayed swallow initiation and weak swallow in post-stroke dysphagia patients to reduce the latency of onset of the pharyngeal phase and its related complications of premature spillage and aspiration. Thus, a possible MI-SW-based or MI-Ton-based swallowing rehabilitation model could be coupled with a suitable end effector such as a visual image and tactile feedback with vibration or subtetanic electrical stimulation in the submental area. This technique would need to be closely combined with conventional dysphagia rehabilitation in the same treatment episode due to the complexity of swallowing.

Acknowledgments

The authors would like to thank the anonymous reviewers for their constructive comments that greatly contributed to improvements in the quality of the paper.

References

[1] McKeown M J, Torpey D C and Gehm W C 2002 Non-invasive monitoring of functionally distinct muscle activations during swallowing *Clin. Neurophysiol.* **113** 354–66

[2] Kiger M, Brown C S and Watkins L 2006 Dysphagia management: an analysis of patient outcomes Using VitalStim^{TM} therapy compared to traditional swallow therapy *Dysphagia* **21** 243–53

[3] Fraser C, Power M, Hamdy S, Rothwell J, Hobday D, Hollander I, Tyrell P, Hobson A, Williams S and Thompson D 2002 Driving plasticity in human adult motor cortex is associated with improved motor function after brain injury *Neuron* **34** 831–40

[4] Dziewas R, Soros P, Ishii R, Chau W, Henningsen H, Ringelstein E B, Knecht S and Panteva C 2003 Neuroimaging evidence for cortical involvement in the preparation and in the act of swallowing *Neuroimage* **20** 135–44

[5] Aboofazeli M and Moussavi Z 2006 Analysis of temporal pattern of swallowing mechanism *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 5591–4

[6] Martin R E, MacIntosh B J, Smith R C, Barr A M, Stevens T K, Gati J S and Menon R S 2004 Cerebral areas processing swallowing and tongue movement are overlapping but distinct: a functional magnetic resonance imaging study *J. Neurophysiol.* **92** 2428–43

[7] Sharma N, Pomeroy V M and Baron J-C 2006 Motor imagery: a backdoor to the motor system after stroke? *Stroke* **37** 1941–52

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

[9] Ang K K, Guan C, Chua K S G, Ang B T, Kuah C W K, Wang C, Phua K S, Chin Z Y and Zhang H 2011 A large clinical study on the ability of stroke patients in using EEG-based motor imagery brain-computer interface *Clin. EEG Neurosci.* **42** 253–8

[10] Page S J, Levine P, Sisto S A and Johnston M V 2001 Mental practice combined with physical practice for upper-limb motor deficit in subacute stroke *Phys. Ther.* **81** 1455–62

[11] Dickstein R, Dunsky A and Marcovitz E 2004 Motor imagery for gait rehabilitation in post-stroke hemiparesis *Phys. Ther.* **84** 1167–77

[12] Furlong P L, Hobson A R, Aziz Q, Barnes G R, Singh K D, Hillebrand A, Thompson D G and Hamdy S 2004 Dissociating the spatio-temporal characteristics of cortical neuronal activity associated with human volitional swallowing in the healthy adult brain *Neuroimage* **22** 1447–55

[13] Yang H, Guan C, Ang K K, Wang C, Phua K S and Yu J 2012 Dynamic initiation and dual-tree complex wavelet feature-based classification of motor imagery of swallow EEG signals *Proc. Int. Jt. Conf. Neural Netw.* 1–6

[14] Yang H, Guan C, Ang K K, Wang C, Phua K S, Tang K Y and Zhou L 2013 Feature consistency-based model adaptation in session-to-session classification: A study using motor imagery of swallow EEG signals *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 429–32

[15] Yang H, Guan C, Ang K K and Wang C 2013 Detection of motor imagery of swallow with model adaptation: swallow or tongue? *Abstract in Fifth Int. Brain Computer Interface Meeting* 56

[16] NeuroScan NuAmps http://www.neuroscan.com

[17] Kingsbury N 1997 Image processing with complex wavelets *Phil. Trans. R. Soc. Lond.* A **357** 2543–60

[18] Selesnick I W 2004 The double-density dual-tree DWT *IEEE Trans. Signal Process.* **52** 1304–14

[19] Selesnick I W, Baraniuk R G and Kingsbury N G 2005 The dual-tree complex wavelet transform *IEEE Signal Process. Mag.* **22** 123–51

[20] Pfurtscheller G and Lopes da Silva F H 1999 Event-related EEG/MEG synchronization and de-synchronization: basic principles *Clin. Neurophysiol.* **110** 1842–57

[21] Pan S J and Yang Q 2010 A survey on transfer learning *IEEE Trans. Knowl. Data Eng.* **22** 1345–59

[22] Li Y, Kambara H, Koike Y and Sugiyama M 2010 Applications of covariate shift adaptation techniques in brain-computer interfaces *IEEE Trans. Biomed. Eng.* **57** 1318–24

[23] Vidaurre C, Kawanabe M, von Bunau P, Blankertz B and Muller K R 2011 Toward unsupervised adaptation of LDA for brain-computer interfaces *IEEE Trans. Biomed. Eng.* **58** 587–97

[24] Shenoy P, Krauledat M, Blankertz B, Rao R P and Muller K R 2006 Towards adaptive classification for BCI *J. Neural Eng.* **3** R13–23

[25] Dai W, Yang Q, Xue G-R and Yu Y 2007 Boosting for transfer learning *Proc. Int. Conf. Mach. Learn.* 193–200

[26] Jain A K, Duin Robert P W and Mao J 2000 Statistical pattern recognition: a review *IEEE Trans. Pattern Anal. Mach. Intell.* **22** 4–37

[27] Muller K, Mika S, Ratsch G, Tsuda K and Scholkopf B 2001 An introduction to kernel-based learning algorithms *IEEE Trans. Neural Netw.* **12** 181–201
[28] Chang C C and Lin C J 2011 LIBSVM: a library for support vector machines *ACM Trans. Intel. Syst. Tech.* **2** 1–27 Software available at http://csie.ntu.edu.tw/˜cjlin/libsvm
[29] Blankertz B, Tomioka R, Lemm S, Kawanabe M and Muller K R 2008 Optimizing spatial filters for robust EEG single-trial analysis *IEEE Signal Process. Mag.* **25** 41–56
[30] Ramoser H, Muller-Gerking J and Pfurtscheller G 2000 Optimal spatial filtering of single trial EEG during imagined hand movements *IEEE Trans. Rehabil. Eng.* **8** 441–6
[31] 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
[32] Sun G, Hu J and Wu G 2010 A novel frequency band selection method for common spatial pattern in motor imagery based brain computer interface *Proc. Int. Jt. Conf. Neural Netw.* 1–6
[33] Teismann I K, Warnecke T, Suntrup S, Steinstrater O, Kronenberg L, Ringelstein E B, Dengler R, Petri S, Pantev C and Dziewas R 2011 Cortical processing of swallowing in ALS patients with progressive dysphagia-a magnetoencephalographic study *PLoS ONE* **6** e19987
[34] Croft R J and Barry R J 2000 Removal of ocular artifact from the EEG: a review *Neurophysiol. Clin.* **30** 5–19
[35] De Clercq W, Vergult A, Vanrumste B, Van Paesschen W and Van Huffel S 2006 Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram *IEEE Trans. Biomed. Eng.* **53** 2583–7
[36] Gomez-Herrero G, De Clercq W, Anwar H, Kara O, Egiazarian K, Van Huffel S and Van Paesschen W 2006 Automatic removal of ocular artifacts in the eeg without a reference eog channel *Proc. Nordic Signal Process. Symp.* 130–3