Journal of Neuroscience Methods xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Computational Neuroscience

Detection of motor imagery of brisk walking from electroencephalogram

Huijuan Yang*, Cuntai Guan, Chuan Chu Wang, Kai Keng Ang

*Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore*

HIGHLIGHTS

- Proposed a novel method for detection of motor imagery of walking from background idle state.
- Proposed an optimization/regularization method to jointly select both channels and frequency bands simultaneously.
- In selecting channels and frequency bands, we consider: dependency between features and class labels, redundancy between to-be selected with selected features, and separation between classes.

ARTICLE INFO

Article history:
Received 15 January 2014
Received in revised form 3 May 2014
Accepted 6 May 2014
Available online xxx

Keywords:
Rehabilitation
Motor imagery of walking
Optimization/regularization
Maximum dependency and minimum redundancy
Class separation
Joint channel and frequency band selection

ABSTRACT

Rehabilitation of lower limbs is equally as important as that of upper limbs. This paper presented a study to detect motor imagery of walking (MI-Walking) from background idle state. Broad overlapping neuronal networks involved in reorganization following motor imagery introduce redundancy. We hypothesized that MI-Walking could be robustly detected by constraining dependency among selected features and class separations. Hence, we proposed to jointly select channels and frequency bands involved in MI-Walking by optimizing/regularizing the objective function formulated on the dependency between features and class labels, redundancy between to-be-selected with selected features, and separations between classes, namely, "regularized maximum dependency with minimum redundancy-based joint channel and frequency band selection (RMDR-JCFS)". Evaluated on electroencephalography (EEG) data of 11 healthy subjects, the results showed that the selected channels were mainly located at premotor cortex, mid-central area overlaying supplementary motor area (SMA), prefrontal cortex, foot area sensory cortex and leg and arm sensorimotor representation area. Broad frequencies of alpha, mu and beta rhythms were involved. Our proposed method yielded an averaged accuracy of 76.67%, which was 9.08%, 5.03%, 7.03%, 14.15% and 3.88% higher than that obtained by common spatial pattern (CSP), filter-bank CSP, sliding window discriminate CSP, filter-bank power and maximum dependency and minimum redundancy methods, respectively. Further, our method yielded significantly superior performance compared with other channel selection methods, and it yielded an averaged session-to-session accuracy of 70.14%. These results demonstrated the potentials of detecting MI-Walking using proposed method for stroke rehabilitation.

1. Introduction

Stroke is one of the leading causes of mortality and disability of adults in most industrialized countries (Dunsky et al., 2008; Diaz et al., 2011). How to design training tools to help the patients achieve independence in activities of daily living is a major concern in rehabilitation. Restoring walking ability or improving gait is one of the major concerns in stroke rehabilitation, since roughly one third of surviving patients had lost their independent walking ability or walking in a slow and asymmetric manner (Dunsky et al., 2008). Despite the encouraging results achieved in using motor imagery (MI)-based training tools for upper extremity such as hand and arm rehabilitation, the use of motor imagery for lower limb rehabilitation only received sufficient attention recently (Dickstein et al., 2004; Castermans et al., 2014). Motor imagery of movements may be a good alternative to activate the neuronal circuits involved in movements since the hemiparesis of stroke patients

* Corresponding author. Tel.: +65 64082708.
E-mail addresses: hjyang@i2r.a-star.edu.sg (H. Yang), ctguan@i2r.a-star.edu.sg (C. Guan), ccwang@i2r.a-star.edu.sg (C.C. Wang), kkang@i2r.a-star.edu.sg (K.K. Ang).

http://dx.doi.org/10.1016/j.jneumeth.2014.05.007
0165-0270/

had prevented the physical movements of their limbs (Dickstein et al., 2004). The activity-dependent plasticity throughout center nervous system influences the functional outcome of the patients. Mental imagery caused reorganization of the functional networks in both healthy controls and stroke patients (Dunsky et al., 2008). Clinical studies consistently found improved motor performance by combining motor imagery with physical therapies, compared with those using physical or occupational therapies in separation. The efficacy of using motor imagery for rehabilitation was established in acute, chronic, mild and severe hemiparesis (Dickstein and Deutsch, 2007; Ang et al., 2011). It was anticipated that the rhythmic foot or leg movements activated the primary motor cortex, while the movement preparation activated the frontal and associated areas (Diaz et al., 2011; Dunsky et al., 2008).

Conventional techniques for gait rehabilitation acted on distal physical level but with the aim to influence the top neurosystem (Belda-Lois et al., 2011). Integrating different strategies in neurophysiological and motor learning techniques appeared to be more effective compared with using one single strategy alone. Combination of functional electrical stimulation (FES) with different walking retraining strategies improved the hemiplegic gait with faster rehabilitation and improved endurance (Belda-Lois et al., 2011). Robotics has emerged as a rehabilitation treatment tool to provide safe, intensive, repetitive and task-oriented trainings, and controllable assistance. This has addressed the intensive labors involved in traditional rehabilitation therapies for gait (Diaz et al., 2011). Furthermore, quantifiable measures of subject performance and interactive training through bio-feedback are possible by using robotic devices for both upper and lower limbs training. Non-invasive gait training at home based on MI-Walking justified its feasibility for home-based lower limb training (Dickstein et al., 2004; Dunsky et al., 2008). Motor imagery-based training allowed intensive and repetitive motion to be performed mentally at home with reasonable cost, and allows quantitative assessment of level of recovery (Dunsky et al., 2008). It could even be used for patients with no residual motor function (Diaz et al., 2011). In one study, participants of unilateral stroke received task-specific training for 6 weeks (Dunsky et al., 2008). Improvements were observed in walking speed, stride length, cadence and single-support time. Whereas in another single case study, improvements in walking speed were noticed with the gains partially extended beyond the practice period (Dickstein et al., 2004). Motor imagery generates event-related (de)synchronization (ERD/ERS). Beta-rebound reflected the active inhibition of neuronal networks after termination of a sensorimotor program (Solis-Escalante et al., 2010, 2012). The peri-imagery ERD and post-imagery ERS were utilized to realize a brain switch (Pfurtscheller and Solis-Escalante, 2009; Solis-Escalante et al., 2010; Muller-Putz et al., 2010). One single Laplacian derivatives and a full description of frequency band powers were used to detect the brisk foot movements execution (Solis-Escalante et al., 2008). Other works on detection of lower limb movement and imagination include: detection of dorsiflexion of both feet (Solis-Escalante et al., 2008), comparisons of the effects for motor imagery of foot dorsiflexion and gait based on motor evoked potentials, and application of transcranial magnetic stimulation over the primary motor cortex (Bakker et al., 2008). Combining observation with motor imagery enhanced activation compared with that of observation only (Villiger et al., 2013), which motivated us to use avatar walking as a cue in the experiments.

Recent developments of non-invasive BCIs dedicated to motor rehabilitation was reviewed in Castermans et al. (2014), which focused on principles of human locomotion control and mechanisms of supra-spinal centers being active during gait. Neuro-physiological signals such as EEG and upper limb EMG were used to control assistive exoskeletons in locomotion (Cheron et al., 2012). Classification of the EEG signals of repetitive foot dorsiflexion (Do et al., 2011) or walking kinesthetic motor imagery (KMI) (Do et al., 2013) from idling was used to trigger FES of tibialis anterior muscle of the contralateral foot (Do et al., 2011), or the robotic gait orthosis (Do et al., 2013). Beta oscillations generated by MI of the paralyzed feet were used to control the wheelchair (Wang et al., 2012). Better accuracy was achieved in detecting movements from EEG in a self-paced asynchronous BCI using the movement-related cortical potentials (MRCPs) with decreased accuracy for MI (Niazi et al., 2011). Single-trial classification of gait intent from a point intent or standing in place using regularized linear discriminant analysis based on principal components analysis (PCA) reduced feature space was investigated (Velu and de Sa, 2013). High accuracies in predicting movements execution (ME) or MI using EEG signals preceding the ME/MI from sensorimotor areas were associated with classifications that relied heavily on the ERD/ERS (Morash et al., 2008). Coupling of the electrocortical activity with gait cycle phase during steady-speed human walking was supported by the significant spectral power increase in alpha and beta band at the sensorimotor motor cortex (Gwin et al., 2011).

Channel selection is important due to the following reasons. Firstly, it is important to identify the brain regions relating to motor imagery tasks performed (Lal et al., 2004), especially for stroke patients with lesions in motor cortex (Tam et al., 2011). Secondly, channel selection is necessary to avoid over-fitting of classifiers and spatial filters with the increased number of irrelevant channels. Thirdly, reducing the number of channels would significantly reduce the prolonged setup time and cost of system. Channel selection can be done by utilizing existing feature selection methods on the paradigm of motor imagery, which can be classified as “wrapper-based method”, e.g., coupled with a classifier (Lal et al., 2004; Schroder et al., 2005; Tam et al., 2011), or “filter-based method”, e.g., optimization with some criterion (Lan et al., 2005; Wang et al., 2005; Farquhar et al., 2006; Yong et al., 2008; Arvaneh et al., 2011). Channels with the maximum coefficients of spatial patterns were considered as the most correlated to corresponding task in CSP-based channel selection (cCSP) (Wang et al., 2005). Maximizing the mutual information (MI) between features and class labels was considered in MI-based channel selection (cMI) (Lan et al., 2005). Regularizing the spatial filter to minimize the number of non-zero entries of the weight vector was carried out in regularizing CSP-based channel selection (cCSP) (Yong et al., 2008; Farquhar et al., 2006). Sparsifying the common spatial filters by constraining the classification accuracy was considered in Sparse-CSP based channel selection (csCSP) (Arvaneh et al., 2011). Recursive features elimination (cRFE) and zero-norm optimization (cZNO) based on support vector machines (SVMs) selected the subject-specific channels by eliminating the channels with the smallest distances to the margins or with the smallest mean (Lal et al., 2004). Fisher criterion-based channel selection (cFC) (Lal et al., 2004) determined the correlation of a feature with the labels by calculating the fisher score.

Existing approaches selected channels based on the EEG signals filtered at pre-selected multiple frequency bands. The joint relationship between the frequency band and channels, and their simultaneous influences on classification accuracies have not been explored. Based on the EEG data of MI-Walking collected from 11 healthy subjects, the objective of this study was to investigate how to robustly detect MI-Walking from background idle state by jointly selecting the channels and frequency bands that contributed mostly to the mental task. To achieve this objective, we proposed to jointly select the channels and frequency bands so that the dependency between the power features of a channel and frequency band with the class labels, the redundancy between the to-be-selected feature with those already selected ones, and the class separations can be jointly optimized/regularized.

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007

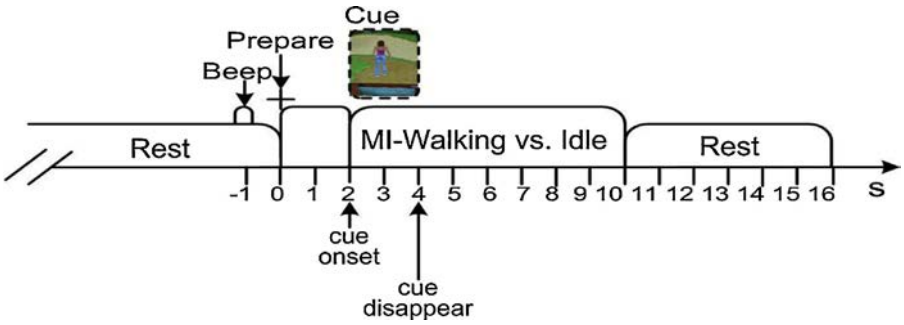

Fig. 1. Timing schemes of the experimental paradigm.

2. Methods

2.1. EEG data collection

EEG data were collected from 11 healthy subjects with ages varying from 24 to 47 years old. The average ages of the subjects were 34.27 ± 8.43 years. Among the subjects, four were females and seven were males. Written informed consents were obtained from the subjects, and ethical approval was obtained from the institutional review board prior to experiments. None of the subjects had the history of neurological or orthopedic disorders. The experimental paradigm was illustrated in Fig. 1. One trial was consisted of three stages of *preparation*, *task performing* (e.g., action or idle) and *resting*. The *preparation* cue shown as changing of traffic lights lasted for 2 s. An acoustic sound 'beep' was played at the end of preparation to remind the subject to start performing the tasks. An avatar character in walking or stance state was used as the cue for *action* or *idle*. The action cue lasted for 2 s. Following the disappearance of the cue, the subject started performing MI-Walking at the comfortable pace, or did nothing by just looking at the screen for 6 s. This was followed by a resting period of 6 s plus a random timing of 0–1.25 s between any two trials. In MI-Walking, the subjects were asked to imagine walking using the two legs with the focus on the rhythmic movements of the legs and joints, and the feelings when the feet touched and pushed the ground. Two sessions of data were collected on two separate days. Each session contained two runs with each run consisting of 40 trials of MI-Walking and 40 trials of idle. The sequences of two actions being shown were randomized. The devices for data collection were the EEG cap and amplifier of Neuroscan NuAmps, and the acquisition software. The EEG signal was digitally subsampled at 250 Hz, with a resolution of 22 bits and voltage ranges of ± 130 mV. The placements of the electrodes followed the international 10–20 system. Thirty-two channels, i.e., Fp1, Fp2, F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, 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 EEG signals at channels A1 and A2 was used as the reference.

How to evaluate the vividness of MI is important since it is the internal rehearsal without motor output (Cui et al., 2007; Malouin et al., 2007). In our experiments, the whole process was monitored and visually inspected. A clear instruction was given to the subject prior to the experiments on how to imagine walking using the two legs, e.g., try to feel the rhythms of the foot stride and swing of the arm and body, and the feelings when the feet touched and left the ground. Furthermore, a walking avatar was shown in the computer screen for the subject to follow easily. The vividness of motor imagery was post-checked with the subject after each run (Cui et al., 2007; Malouin et al., 2007). A proper instruction was re-enforced to ensure that the subject performed the motor imagery task as instructed.

2.2. Feature extraction and mutual information calculation

Selecting the subject-specific channels was especially important for stroke patients whose motor imagery were not symmetrical for the unaffected and affected sides (Malouin et al., 2008). A total of 9 frequency bands from 4 Hz to 36 Hz covering theta, mu (alpha), beta and low gamma frequency rhythms were employed to filter the EEG signal, which were generally employed for the detection of motor imagery of limb movements (Ramoser et al., 2000; Ang et al., 2011; Arvaneh et al., 2011; Blankertz et al., 2008; Lotte and Guan, 2011). These channels and frequency bands formed a 2D matrix, where each feature was a point in it. Features were subsequently selected from this matrix by jointly considering the channels and frequency bands based on our proposed method to classify MI-Walking from background idle state. Let's firstly denote the EEG signal as: $s(m, c, r)$, where $m = 1, 2, \ldots, n_s$, $c = 1, 2, \ldots, n_c$ and $r = 1, 2, \ldots, n_r$ represented the indexes of samples, channels and trials, respectively. The signal was firstly divided into n_f ($n_f = 9$) non-overlapping frequency bands with the band width of 4 Hz, i.e., [4 8] Hz, [8 12] Hz,. . ., [32 36] Hz. The signal was then filtered by Chebyshev digital filter to obtain the filtered signal s_f. The band power for the kth trial, at cth channel and fth frequency band (denoted as $P_w(k, c, f)$) was calculated by

$$P_w(k, c, f) = 10 \log 10 \left(\sum_{m=1}^{n_s} s_f(k, c, m) s_f(k, c, m) \right) \quad (1)$$

where $k = 1, 2, \ldots, n_r$, $c = 1, 2, \ldots, n_c$, $m = 1, 2, \ldots, n_s$ and $f = 1, 2, \ldots, n_f$ denoted the indexes of trials, channels, samples and frequency bands, respectively.

To jointly select the most informative channels and frequency bands, i.e., the most informative features, the dependency between the features with class labels, and the redundancy between the to-be selected features with those already selected ones were calculated. The dependency between the features of a channel and frequency band (denoted as $w(c, f)$, where w is the brevity of P_w in Eq. (1)) and the class labels was evaluated by the mutual information between them (denoted as $\mathcal{I}(\mathcal{W}(c, f); \mathcal{Y})$), which was given by

$$\mathcal{I}(\mathcal{W}(c, f); \mathcal{Y}) = \sum_{w \in \mathcal{W}} \sum_{y \in \mathcal{Y}} p(w(c, f), y) \log \frac{p(w(c, f), y)}{p(w(c, f)) p(y)} \quad (2)$$

where $\mathcal{W}(c, f)$ and $\mathcal{Y}$ represented the collective set of $w(c, f)$ and y; $p(w, y)$ represented the multivariate density. Generally, the probability density function of the features was difficult to obtain, hence, the membership of features for all the problem classes was obtained by the methods such as fuzzy K-Nearest Neighbor (Chen et al., 1999) and fuzzy c-means (Lee et al., 1986). In our proposed method, the fuzzy entropy-based membership computation was adopted

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007

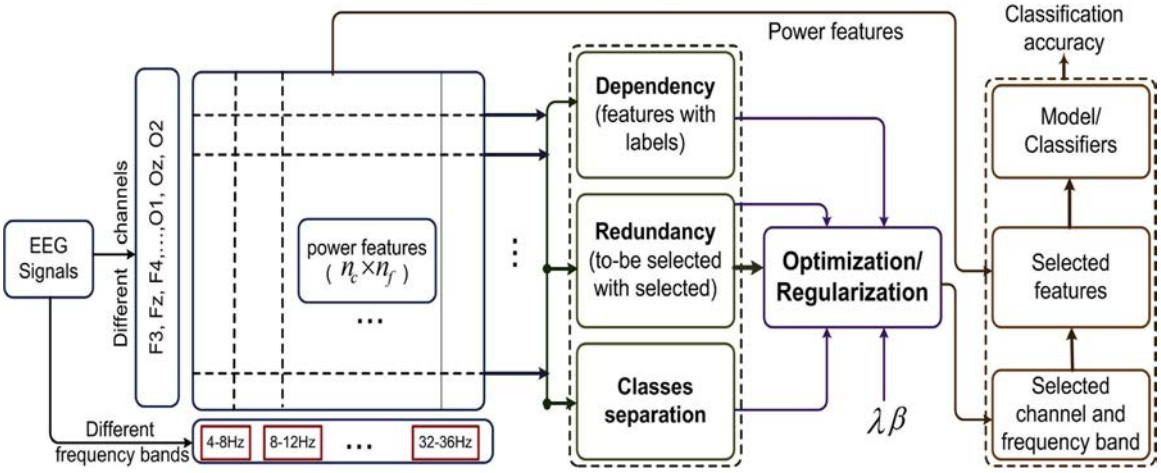

Fig. 2. Schematic illustration of the proposed *regularized maximum dependency with minimum redundancy-based joint channel and frequency band selection* (RMDR-JCFS) method.

(Khushaba et al., 2011). The inter (intra)-class distance (denoted as $\mathcal{D}_{ij}$) was given by

$$\mathcal{D}_{ij} = \left(\frac{||\bar{w}(i) - w(j,k)||_\delta}{r+\epsilon}\right)^{-2/(f_z-1)} \quad (3)$$

where f_z was the fuzzification parameter and $\epsilon > 0$ was used to avoid singularity. δ was the standard deviation of data. $\bar{w}(i)$ and $w(j,k)$ represented the mean feature of class i and kth feature of class j, respectively. $||\,.\,||$ denoted l_2 norm. $\mathcal{D}_{ij}$ represented the intra-class distances when $i=j$; whereas it represented the inter-class distances when $i \neq j$. r denoted the radius of the data, which was given by

$$r = \max||\bar{w}(i) - w(j,k)||_\delta \quad (4)$$

Finally, the membership was assigned such that $\sum_{i=1}^{y_c} \mathcal{D}_{ik} = 1$, where y_c was the total number of classes. The distance between each feature vector and that of the class center was firstly calculated based on a distance function such as Gaussian. The membership was calculated based on Eq. (3), which was then used to calculate the mutual information (denoted as $\mathcal{I}(\mathcal{W}^g, \mathcal{Y})$) using Eq. (2), $p(w(c,f), y)$ (abbreviated as p_{wy}) was given by

$$C_{wy}(c,f) = \frac{1}{n_r}(\mathcal{W}^g(c,f).*\mathcal{Y}) * (\mathcal{W}^g(c,f).*\mathcal{Y})^T \quad (5)$$

$$P_{wy}(c,f) = C_{wy}|C_{wy}(c,f) > 0 \quad (6)$$

where $\mathcal{W}^g(c,f)$ and n_r denoted membership of features (i.e., *grouped* features) at channel c and frequency band f, and total number of trials, respectively. The equation used to compute mutual information between to-be selected with selected features was similar to Eqs. (5) and (6), however, the labels should be replaced with features of those already selected features.

2.3. Proposed joint channel and frequency band selection

The power features of a channel and frequency band that were mostly correlated with the ERD/ERS generated during motor imagery tasks were different from subject to subject, which also varied for different mental tasks. To extract the most discriminate features that were jointly determined by the channels and frequency bands for the detection of MI-Walking from background idling state, we proposed to select the features based on the *maximum dependency with minimum redundancy* criterion as inspired by the feature selection method (Peng et al., 2005; Yang et al., 2013). However, by imposing the constrains on the dependency of features with class labels (DFLs), and the redundancy between to-be-selected feature and those already selected ones (RFFs), the separability of classes (SCCs) cannot be guaranteed. Moreover, the jointly selected frequency band and channels would appear to be unstable, these eventually would lead to the degradation of the performance, especially for session-to-session classification. Hence, in this paper, we proposed a regularization and optimization-based approach to select the features such that the DFLs, RFFs and SCCs were better balanced, namely, "*Regularized Maximum Dependency with Minimum Redundancy-based Joint Channel and Frequency Band Selection* (RMDR-JCFS)", as illustrated in Fig. 2. Conventionally, the top n_d features that maximized $\mathcal{I}(\mathcal{W}(c,f); \mathcal{Y})$ were selected in *maximum dependency* (MD)-based feature selection by

$$(\hat{c}, \hat{f}) = \underset{c\in\{1,2,\ldots,n_c\}, f\in\{1,2,\ldots,n_f\}}{\arg\max} \mathcal{I}(\mathcal{W}^g(c,f); \mathcal{Y}) \quad (7)$$

Generally, the features can be selected by considering the MI between each to-be-selected feature of all the trials with class labels, or selecting a group of features by considering the MI between each collective subset of features with the class labels. Selecting a subset of n_d features from a total of n_t features requires $\binom{n_t}{n_d}$ computations, which would be high when n_t is large. Further, selecting each feature individually without considering the dependency between the features would introduce redundancy, which eventually would degrade the discriminability of features. The constraints of the minimum redundancy of the to-be-selected features with those already selected ones were therefore imposed in the proposed method.

The mutual information between the *grouped* power features with class labels (denoted as $\mathcal{I}_{wy}(c,f)$), and the mutual information (dependency) between the *grouped*, to-be-selected (e.g., jth) feature with those already selected ones (e.g., ith feature) (denoted as $\mathcal{I}_{ww}(c,f)$) for the considering channel (c) and frequency band (f) were given by

$$I_{wy}(c,f) = \underbrace{\hat{I}(\mathcal{W}_j^g(c,f); Y)}_{\mathcal{W}_j(c,f)\in\{\mathcal{W}-\mathcal{W}_{m-1}\}} \quad (8)$$

$$I_{ww}(c,f) = \frac{1}{m-1}\sum_{i=1}^{m-1} \underbrace{\hat{I}(W_j^g(c,f); W_i^g(\widetilde{c},\widetilde{f}))}_{W_i(\widetilde{c},\widetilde{f})\in\mathcal{W}_{m-1}} \quad (9)$$

where $\mathcal{W}_i^g(\widetilde{c},\widetilde{f})$ and $\mathcal{W}_j^g(c,f)$ denoted the ith and jth *grouped* power features at channel $\widetilde{c}$ and frequency band $\widetilde{f}$ in the selected set $\mathcal{W}_{m-1}$ and unselected set $\mathcal{W}-\mathcal{W}_{m-1}$, respectively. $i=1, 2, \ldots, m$ denoted the selected features. To better scale the data, I_{wy} and I_{ww} were normalized by

$$\tilde{I}_{wy} = I_{wy}/\max(abs(I_{wy})) \tag{10}$$

$$\tilde{I}_{ww} = I_{ww}/\max(abs(I_{ww})) \tag{11}$$

Obviously, $\{\tilde{I}_{wy}, \tilde{I}_{ww}\} \in [-1 \quad 1]$, where $\max(X)$ and $abs(x)$ gave the maximum value in set X and the absolute value of x, respectively.

2.4. Regularized maximum dependency with minimum redundancy-based joint channel and frequency band selection(RMDR-JCFS)

Now let us explore how the most discriminate channels and frequency bands were selected by our proposed method. The distance between each feature and that of the class center was firstly calculated using Eq. (3). The separability between the features of the two classes at channel c and frequency band f (denoted as $S_s(c,f)$) was then calculated by

$$\mathcal{I}_{ss}(c,f) = \mathrm{diag}\left(\frac{\sum_{i=\{0,1\}} C_{wii}(c,f)}{\sum_{i,j\in\{0,1\},j\neq i} C_{bij}(c,f)}\right) \tag{12}$$

where $C_{wii}(c,f)$ and $C_{bij}(c,f)$ ($i, j \in \{0, 1\}$, $i \neq j$) represented the covariance matrices of the within-class distances ($\mathcal{D}_{ii}(c,f)$) and between-classes distances ($\mathcal{D}_{ij}(c,f)$, $i \neq j$) for channel c and frequency band f based on fuzzy membership calculation, which were given by

$$\mathcal{C}_{wii}(c,f) = \mathcal{D}_{ii}(c,f)\mathcal{D}_{ii}^T(c,f) \tag{13}$$

$$\mathcal{C}_{bij}(c,f) = \mathcal{D}_{ij}(c,f)\mathcal{D}_{ij}^T(c,f) \tag{14}$$

$\mathrm{diag}(X)$ returns the diagonal values of matrix X.

Let the ith row vector in matrix $\tilde{I}_{wy}$, $\tilde{I}_{ww}$ and I_{ss} be denoted as $\mathbf{I}_{wy}^i$, $\mathbf{I}_{ww}^i$ and $\mathbf{I}_{ss}^i$, respectively, where $i=1, 2, \ldots, n_c$. There are n_c row vectors and each vector is of length n_f, the joint channel and frequency band selection was then formulated as an optimization/regularization problem as

$$\begin{aligned} &\text{minimize} \quad J, \\ &\text{subject to} \quad -1 \le \tilde{I}_{wy} \le 1 \& -1 \le \tilde{I}_{ww} \le 1, \end{aligned} \tag{15}$$

where the objective function j is defined as

$$J(\tilde{I}_{wy}, \tilde{I}_{ww}, \tilde{I}_{ss}, \lambda, \beta) = \lambda\mathcal{R}_{wy} + (1-\lambda)(\beta\mathcal{R}_{ww} + (1-\beta)\mathcal{R}_{ss}) \tag{16}$$

$\mathcal{R}_{wy}$, $\mathcal{R}_{ww}$ and $\mathcal{R}_{ss}$ dealt with dependency between features and class labels, redundancy between to-be-selected with selected features, and separation of the two classes, which were given by

$$\mathcal{R}_{wy} = (1-\mathbf{I}_{wy}^i)(1-\mathbf{I}_{wy}^i)^T \tag{17}$$

$$\mathcal{R}_{ww} = \frac{||\mathbf{I}_{ww}^i(\mathbf{I}_{ww}^i)^T||_1}{||\mathbf{I}_{ww}^i(\mathbf{I}_{ww}^i)^T||_2} \tag{18}$$

$$\mathcal{R}_{ss} = \frac{||\mathbf{I}_{ss}^i(\mathbf{I}_{ss}^i)^T||_1}{||(\mathbf{I}_{ss}^i(\mathbf{I}_{ss}^i)^T)||_2} \tag{19}$$

where $0 \le \lambda \le 1$ was the regularization parameter that controlled the dependency between the features and class labels, and the redundancy between to-be-selected with selected features; and the sparsity of the redundancy and class separation matrices. The balance of the latter two were controlled by parameter β, where $0 \le \beta \le 1$. Ideally, $\tilde{I}_{ww}$ and I_{ss} were expected to be as sparse as possible, i.e., the selected feature should be as less dependent on selected features as possible, and the class separation should be as good as possible. These two terms were penalized with the l_1 over l_2 norms. l_1 and l_2 norms were known to be robust against outliers, and to penalize small residuals while preserving the modest ones, respectively. l_1/l_2 was scale-invariant and had shown good performance in EEG channel selection (Arvaneh et al., 2011). Hence, l_1/l_2 was selected in our method.

The resultant output matrices from the optimization process (denoted as $\tilde{O}_{wy}$, $\tilde{O}_{ww}$ and $\tilde{O}_{ss}$) were then used to jointly select the channel and frequency band, which was formulated as regularization problem as follows

$$(\hat{c},\hat{f}) = \arg\max[\lambda\tilde{O}_{wy}(c,f) + (1-\lambda)\mathcal{R}_o(c,f)] \tag{20}$$

$$\mathcal{R}_o(c,f) = \beta(1-\tilde{O}_{ww}(c,f)) + (1-\beta)(1-\tilde{O}_{ss}(c,f)) \tag{21}$$

The same λ and β employed in the optimization process were used in Eqs. (20) and (21). The optimization problem in Eq. (15) can be seen as a least square or l_2 norm convex quadratic function, which was given by

$$L(x_1,x_2,x_3,\lambda,\beta) = \lambda(1-x1)^2 + (1-\lambda)(\beta x_2^2 + (1-\beta)x_3^2) \tag{22}$$

Eq. (22) can be solved analytically. Indeed, L was a convex function with its Hessian matrix given by

$$H(x_1,x_2,x_3) = \begin{pmatrix} 2\beta & 0 & 0 \\ 0 & 2(1-\beta)\lambda & 0 \\ 0 & 0 & 2(1-\beta)(1-\lambda) \end{pmatrix} \tag{23}$$

Obviously $H(x_1, x_2, x_3)$ was positive definite and the constraint of the problem was a linear function. Hence, finding minimum was possible.

3. Results

Experiments were conducted based on 11 healthy subjects to demonstrate effectiveness of the proposed method. Support vector machine classifiers with linear kernel functions and quadratic programming were selected in the experiments. It should be noted that no outlier detection was performed and no trial was excluded from processing at any stages of the experiments.

3.1. ERD/ERS visualization

We firstly showed the time frequency maps generated using power features of MI-Walking with reference to background idle to visualize the ERD/ERS patterns associated with MI-Walking. The powers of the band-pass filtered EEG signal at channel (c) and frequency band (f) were firstly calculated using Eq. (1), which were then averaged across all the trials of MI-Walking class and idle class to obtain $P_a(c,f)$ and $P_l(c,f)$, respectively. A time segment consisting the *preparation* time of 2 s, *cue* time of 2 s and *action* time of 6 s was used to demonstrate the changes of power energy during the entire time course. $P_a(c,f)$ and $P_l(c,f)$ were then normalized by

$$\tilde{P}_a(c,f) = \frac{P_a(c,f) - R_m(c,f)}{R_x(c,f) - R_m(c,f)} \tag{24}$$

where $R_x(c,f) = \max(\max(P_a(c,f)))$ and $R_m(c,f) = min(min(P_a(c,f)))$ denoted the maximum and minimum of the powers at channel c and frequency band f. $P_l(c,f)$ was similarly normalized to obtain $\tilde{P}_l(c,f)$. The difference between power features of MI-walking and idle at channel c and frequency band f (denoted as $P_d(c,f)$) was computed by $P_d(c,f) = \tilde{P}_l(c,f) - \tilde{P}_a(c,f)$ which was used to plot the time frequency maps, as shown in Fig. 3 for 4 representative subjects examined at electrodes 'C3', 'Cz', 'CPz' and 'Pz'. Performing MI-Walking (with reference to idling) resulted in ERD and ERS in the motor area and medial foot representation area in primary motor cortex at *alpha*, *mu* and *low beta* frequency bands, which was

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007

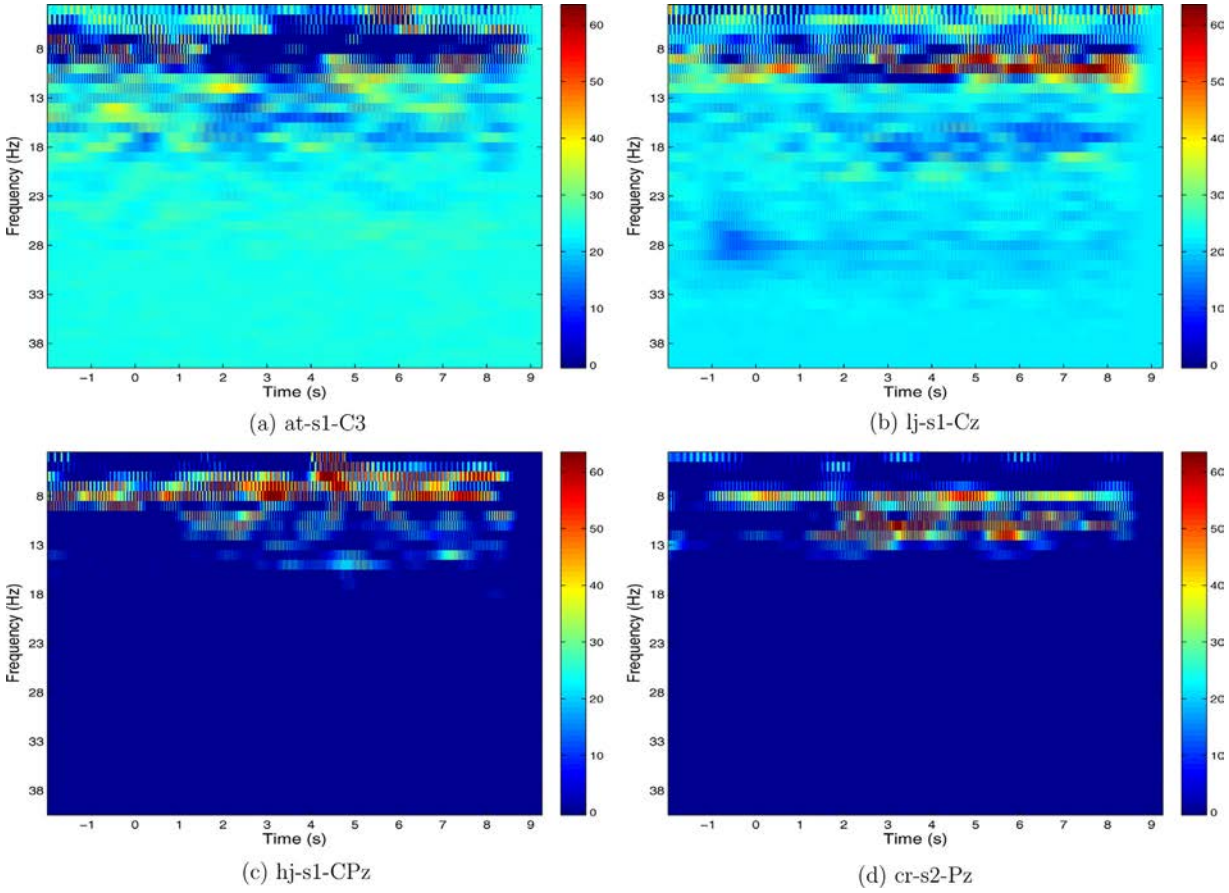

Fig. 3. Examples of ERD/ERS visualization. The time–frequency plots of MI-Walking with reference to background idle state. "at-s1-C3" represented subject 'at' at session 1 and electrode 'C3'; colors depicted the relative powers; the timings of *x*-axis represented: −2 to 0 s (preparation), 0–2 s (action cue), 2 s (action cue disappeared), 2–8 s (MI-Walking/idle) and 8–9 s (post MI-Walking/idle). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

consistent with the hand area ERD (e.g., 'C3') and foot area ERS (e.g., 'Cz', 'CPz', 'Pz') from voluntary foot movements (Pfurtscheller and da Silva, 1999; Severens et al., 2012). These results agreed with the suppression of *mu* and *beta* rhythms during active walking in the lokomat gait orthosis, or active compared to passive walking (Wagner et al., 2012; Severens et al., 2012; Castermans et al., 2014). The post-imagination beta ERS at 'Cz' and 'CPz' (Fig. 3(b) and (c)) was supported by the robust post-movements/imagination ERS after hand and foot movements/imagery (Pfurtscheller and da Silva, 1999; Muller-Putz et al., 2010; Pfurtscheller and Solis-Escalante, 2009).

3.2. Performance of our proposed method

In this section, we presented the classification accuracies of our proposed method, i.e., RMDR-JCFS for both cross-validation (CV) and session-to-session classifications. The sensitivity of the classification performance to the choice of the number of features (i.e., the channel-frequency pairs) was also examined. Two regularization parameters, i.e., λ and β in Eqs. (16), (20) and (21) need to be defined. We investigated two ways to select the parameters: (a) SB-PARA-CV: to select the subject-specific best parameters by employing a 10-folds cross-validation (CV) from a predefined set (e.g., {0.1, 0.2, . . ., 0.9}) (Lotte and Guan, 2011). The parameters that yielded the maximum normalized median accuracy (normalized by variance) were selected. (b) SB-PARA-BPS: to select the pairs of parameters of λ and β from a predefined set, the pair that yielded the best performances considering all the subjects were chosen. The CV accuracies of the proposed method by selecting different numbers of features (i.e., $N_{fs} = 2$, 4, 6 and 8) using parameters selected by SB-PARA-BPS, and by selecting 4 features using subject-specific parameters selected by: SB-PARA-CV1 from set $\{[0.4 : 0.1 : 0.8]\}$, and SB-PARA-CV2 from set $\{[0.1 : 0.1 : 0.9]\}$, were shown in Fig. 4. The results demonstrated that the CV classification accuracies did not vary significantly with the changing of the numbers of selected channel-frequency pairs. The averaged classification accuracies (%) across subjects were: 74.73 ± 1.80, 75.25 ± 1.79, 76.39 ± 1.80 and 76.67 ± 2.03, 75.41 ± 2.54 and 75.92 ± 2.55 by selecting 2, 4, 6 and 8 features using SB-PARA-BPS, and selecting subject-specific parameters using SB-PARA-CV from two different sets, respectively. This was validated by performing a one way analysis of variance (ANOVA) to compare the mean accuracies (%) obtained with different parameters settings. The results showed that the mean accuracies of proposed method (i.e., RMDR-JCFS) with parameters selected by SB-PARA-BPS of 2 features ($M = 74.73$, $SD = 12.21$), 4 features ($M = 75.25$, $SD = 12.70$), 6 features ($M = 76.39$, $SD = 12.42$), and 8 features ($M = 76.67$, $SD = 12.15$); SB-PARA-CV1 of 4 features ($M = 75.41$, $SD = 12.30$); and SB-PARA-CV2 of 4 features ($M = 75.91$, $SD = 2.04$), were not significantly different, $f(5, 126) = 0.078$, $p = 0.996$, where M and SD were acronyms of mean and standard deviation, respectively. However, selecting a larger number of features will significantly increase the computation time, e.g., the computation time required by selecting 8 features was about 9–10 times longer than that required by selecting 2 features. The 95% confidence estimation of the accuracy performed at chance level for each action was approximately 42.5–57.50% based on binomial inverse cumulative distribution function. The results showed that only subject 'th' at session 2 (subject number 22), who

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007

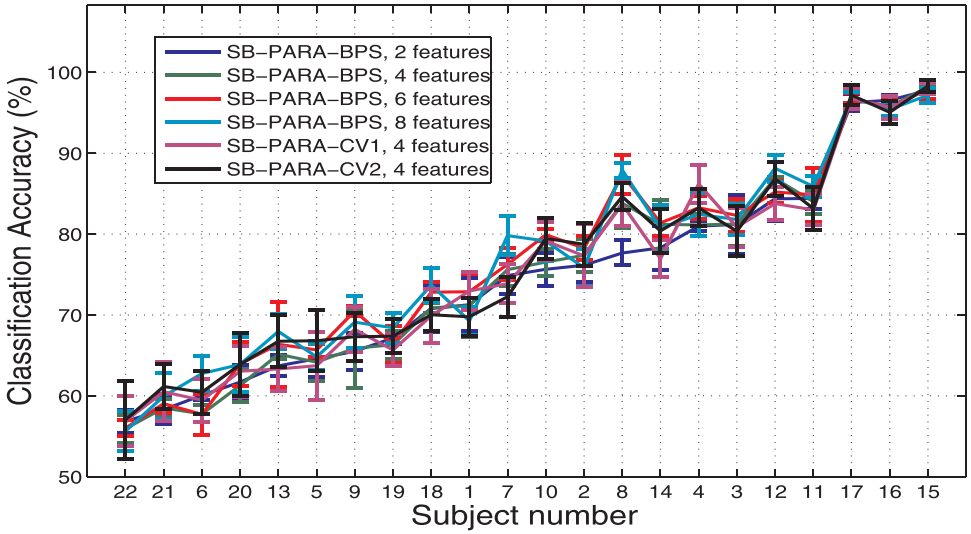

Fig. 4. Classification accuracies of proposed RMDR-JCFS by selecting different numbers of features and subject-specific parameters from different sets.

was BCI-naive, performed at chance level. The session-to-session classification accuracies were evaluated by firstly selecting the features of evaluation session based on the channels and frequency pairs selected by another session, which were then classified by the model generated from another session. The models generated by selecting the subject-specific parameters using SB-PARA-CV from set (e.g., {0.1, 0.2, . . ., 0.9}) was used in evaluation. A moving window of length 2 s was employed with the mean and best accuracies among the windows reported in Table 1. Overall, the session-to-session accuracies were good except subjects 'cc', 'an' and 'th'. The CV accuracies of 'an' and 'th' were very low also. The low session-to-session accuracies of 'cc' may be due to the inconsistencies of the selected channel-frequency pairs across sessions.

The CPU time used in the experiments was reported to show the computational complexity of proposed algorithm. Experiments were conducted using Matlab R2013a, on a desktop computer with Intel Core i7-920 processor of 2.67 GHz, 4 GB RAM and 64-bit windows 7 professional operating system. Note that the

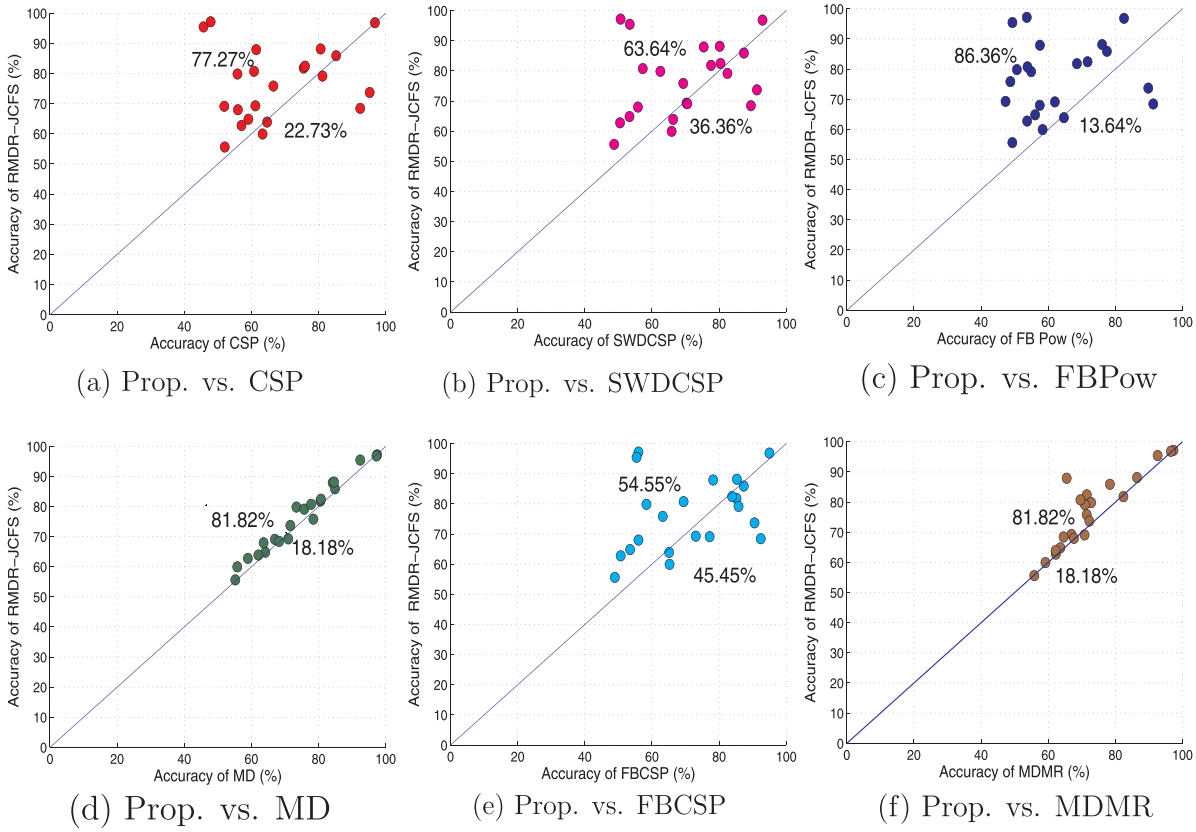

Fig. 5. Pair-wise scatter plot of the accuracies of our proposed method (Prop.) versus other classification methods. In the figure, each dot represented the classification accuracy of a subject. The results demonstrated that our proposed method achieved higher accuracies compared with other methods, i.e., most points lied above the diagonal line.

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007

Table 1
Session-to-session classification accuracies of proposed method.

Sb./Ss.	$(A_a \pm v_r)$	A_b	Sb./Ss.	$(A_a \pm v_r)$	A_b
cc/1	52.08 ± 5.42	55.63	at/1	60.73 ± 15.22	66.88
cc/2	57.81 ± 9.65	63.13	at/2	66.98 ± 53.82	76.25
lj/1	64.90 ± 21.00	72.50	cr/1	89.79 ± 9.48	93.13
lj/2	58.65 ± 34.28	66.25	cr/2	86.77 ± 30.22	91.25
an/1	50.00 ± 0.00	50.00	zm/1	90.73 ± 5.53	92.50
an/2	53.65 ± 9.60	58.13	zm/2	69.79 ± 3.39	72.50
xy/1	65.21 ± 5.26	68.75	mt/1	63.19 ± 7.27	66.25
xy/2	68.85 ± 5.85	73.13	mt/2	61.46 ± 6.04	64.38
ks/1	65.73 ± 8.35	68.75	th/1	50.00 ± 0.00	50.00
ks/2	66.25 ± 9.69	70.63	th/2	51.98 ± 8.97	56.88
hj/1	79.58 ± 9.17	83.13	hj/2	76.88 ± 20.00	83.13
A_{as}	65.95 ± 12.65	A_{bs}	70.14		

Note: $A_a(\%)$, $A_b(\%)$ and v_r: mean and best accuracies and variance. Sb./Ss.: subjects/sessions; A_{as} and A_{bs}: mean and best accuracies across subjects.

following reported time was for a total of 160 trials in each session. The averaged CPU time across subjects for 5×5 CV using the best parameters, i.e., SB-PARA-BPS, were 226.90 ± 3.76 s and 607.43 ± 11.31 s by selecting 4 and 8 features, respectively. By selecting the best parameters using SB-PARA-CV1, the averaged CPU time across subjects for 5×5 CV were 981.33 ± 15.68 s and 2739.60 ± 36.98 s by selecting 4 and 8 features, respectively. These results revealed that selecting regularization parameters and features based on optimization/regularization needed more time, nevertheless, these only required during training phase. The averaged time for session-to-session classification was 57.04 ± 1.37 s for 11 moving windows, i.e. the detection time of each window for each trial was: $0.0324 \pm 7.78\text{e}{-4}$ s.

3.3. Performance comparisons

The performance of our proposed method was compared with existing classification and channel selection methods for EEG signals to demonstrate its effectiveness and advantages. The proposed RMDR-JCFS was firstly compared with several typical classification

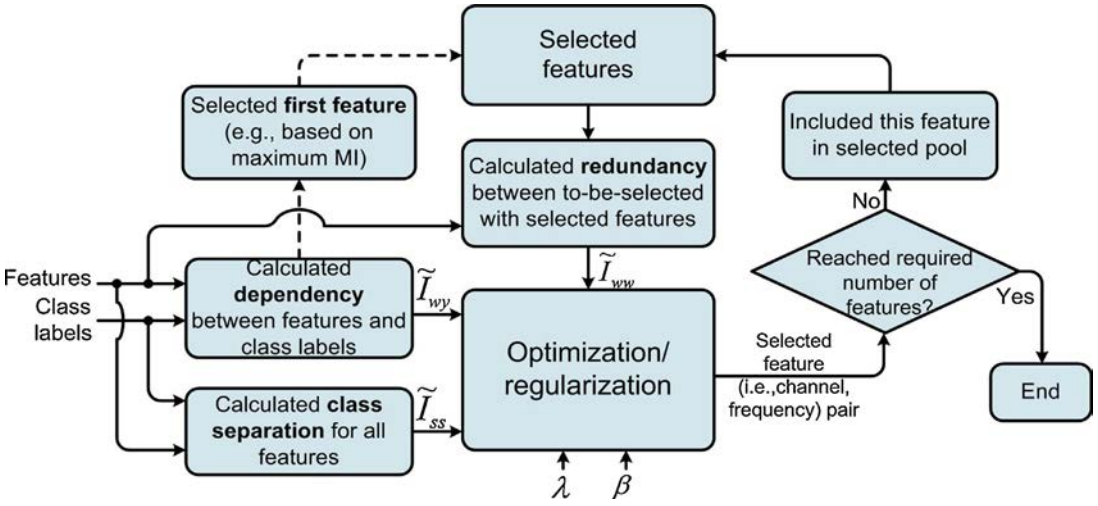

Fig. 6. Schematic illustration of the proposed feature selection process.

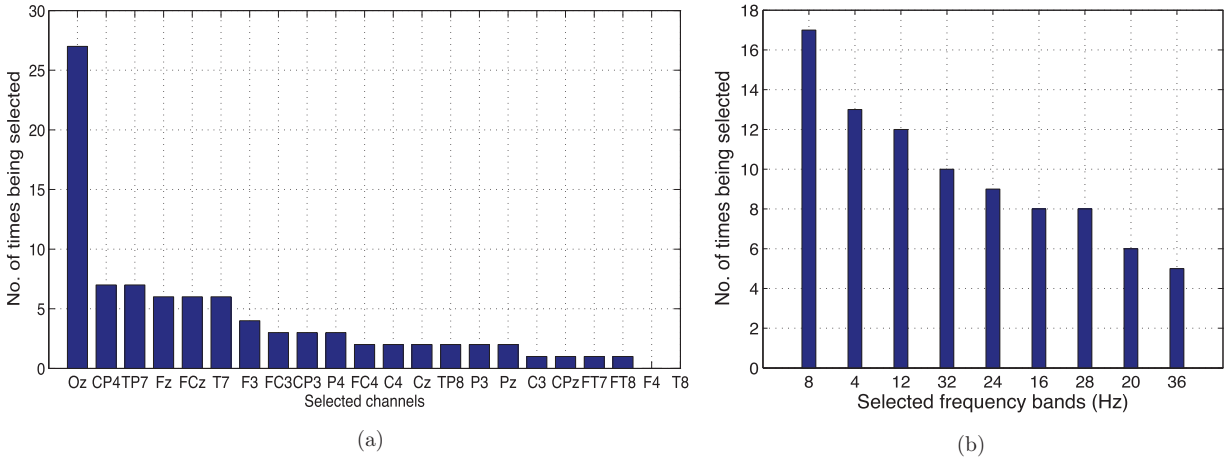

Fig. 7. Overall distributions of selected (a) channels and (b) frequency bands. A total of 4 features were selected, the parameters were selected by SB-PARA-CV2.

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007

Table 2
Performance comparison of classification of motor imagery of walking from idle EEG signals using proposed and other channel selection methods.

Sb./Ss.	*cCSP*	*cMI*	*cFC*	*cZNO*	*cRFE*	*RMDR-JCFS*
cc/1	54.69	68.38	54.94	57.50	65.81	**69.27**
cc/2	61.06	49.88	61.56	49.75	62.63	**75.83**
lj/1	70.19	76.13	69.63	77.06	70.13	**81.81**
lj/2	73.25	74.56	73.50	71.94	75.94	**82.43**
an/1	**65.56**	64.94	63.69	56.13	54.25	64.82
an/2	48.25	55.63	48.63	47.06	51.63	**62.76**
xy/1	52.44	44.00	52.38	51.75	55.25	**79.82**
xy/2	59.56	55.38	59.19	60.50	54.25	**87.91**
ks/1	49.88	49.50	50.56	48.19	58.00	**69.11**
ks/2	76.50	77.06	75.63	74.00	75.31	**79.14**
hj/1	73.44	72.50	74.81	63.69	78.44	**85.89**
hj/2	73.56	70.19	72.06	70.69	74.38	**88.13**
at/1	53.31	55.00	53.19	50.31	48.25	**67.97**
at/2	62.88	61.94	63.44	57.50	60.06	**80.73**
cr/1	94.31	96.50	94.44	93.81	95.63	**97.17**
cr/2	94.00	94.56	93.25	92.38	90.19	**95.44**
zm/1	90.56	87.94	90.44	88.25	87.75	**96.83**
zm/2	70.88	66.38	72.00	64.69	70.69	**73.71**
mt/1	67.63	67.38	67.33	**68.79**	63.71	68.42
mt/2	51.00	54.25	50.00	49.69	51.31	**63.89**
th/1	49.56	51.63	52.44	54.81	52.44	**59.95**
th/2	45.13	52.19	46.13	48.38	49.38	**55.63**
A_{as}	65.35	65.72	65.42	63.49	65.70	**76.67**

Best performances for each subject at each session were shown in bold.

methods for motor imagery EEG signals, which include: filter bank common spatial pattern (FBCSP) (Ang et al., 2012), filter bank with power features (FBPow), common spatial pattern (CSP) (Ramoser et al., 2000; Blankertz et al., 2008), sliding window discriminate CSP (SWDCSP) (Sun et al., 2010), our proposed method by only using the maximum dependency (MD), i.e., λ=1 in Eqs. (16), (20) and (21), and our earlier work, i.e., maximum dependency with minimum redundancy (MDMR) method (Yang et al., 2013). The EEG data from 0.5 s to 2.5 s after onset of the visual cue were used for these CSP-based methods according to the papers. Whereas the optimal time segments obtained by 10-folds cross-validation was employed in our proposed method, MD and MDMR methods. The parameters were selected by SB-PARA-BPS and 8 features (i.e., channel-frequency pairs) were selected for proposed method, whereas $\lambda = 0.5$ was used in MDMR. A pair-wise scatter plot of the 10×10 cross-validation classification accuracies of proposed RMDR-JCFS versus other methods was shown in Fig. 5. The averaged classification accuracies across subjects obtained were: 76.67%, 67.58%, 69.64%, 62.52%, 74.79%, 71.64% and 72.78% for our proposed RMDR-JCFS, CSP, SWDCSP, FBPow, MD, FBCSP and MDMR methods, respectively. These results demonstrated that the averaged classification accuracy obtained by our proposed method was 9.08%, 7.03%, 14.15%, 1.88%, 5.03% and 3.88% higher than that obtained by CSP, SWDCSP, FBPow, MD, FBCSP and MDMR methods, respectively. The superior performance of our proposed method can be clearly seen from Fig. 5, where more points lied above the diagonal

Table 3
Comparison of methods on the detection of motor imagery/movements execution of walking/foot dorsiflexion.

Publication	BCI paradigm subj./device	Methodology	Fre. band results	Activated/used brain areas
Morash et al. (2008)	ME/MI R/L hand squeeze/tongue/R foot toe curl, 8 hea.	ICA filtering, DWT features, FS: BHDT	0–40 Hz, <68.0% (ME), <65.0% (MI)	Motor-related channels
Pfurtscheller and Solis-Escalante (2009)	MI of foot dorsiflexion, 5 hea.	Log band, power peri-ERD and post-ERS	6–36 Hz 59.2 ± 20.3% (asyn.) 78.6 ± 13.5% (cued)	Using LAD of Cz
Muller-Putz et al. (2010)	ME/MI of foot dorsiflexion, 6 hea.	Log band power ME mdl. classify MI	β,69–89% (TPR)	Using LAD of Cz
Niazi et al. (2011)	ME/MI-ballistic ankle dorsiflexions, 19 hea.+5 pat.	MRCP templates, spatial filters, LLSF, OSF, CSP	82.5 ± 7.8% (ME, hea.), 64.5 ± 5.33% (MI, hea.), 55 ± 12.01% (AME, pat.)	0.05–20 Hz multi-chans.
Gwin et al. (2011)	ME-Walking on treadmill, 8 hea./no device	ICA, dipole for ICs, clustering, ERSP, spectrum	α, β, high γ sign. power changes: intra-stride, end stance	ACC posterior parietal sensorimotor cortex
Cheron et al. (2012)	ME-Walking on treadmill, 5 hea./no device	ICA, high-pass filtering, ERSP of C3, C4 chans.	α,β,γ involved in control of walking patterns	Sensorimotor cortex
Wang et al. (2012)	MI-Walking vs. idling, 8 hea. + 1 SCI	Subj.-spec. fre. band PSD features, dim. redu.: CPCA/AIDA	μ,β, 77.2 ± 11% (offline), 8.5 ± 1.1(tt.:10)(online)	Lateral central, centro-parietal (hea.), mid-central (SCI)
Do et al. (2011)	MI-Walking vs. idle,1 hea. + 1 pat. robot gait orthosis	Subj.-spec. fre. band PSD features, dim. redu.: CPCA, AIDA	α, μ, 86.30% (offline), 0.812 ± 0.048 (CCR)	SMA, frontal leg and arm, representation area
Velu and de Sa (2013)	ME-Walk/Point Stand (L/R/F(still)), 9 hea./no device	Wavelet features, dim. redu.: PCA	θ, μ, β, 0.5–56 Hz, 83% (offline)	Leg motor area, arm motor area
Proposed	MI-Walking vs. idle, 11 hea./no device	Joint chan./fre. band selection, 2–8 features	α, θ, β, low γ, 76.67 ± 2.03% (offline) 70.14 (S2S)	SMA, motor cortex, midcentral of motor cortex

Note: Subj.: subject; chan.: channel; fre.: frequency; S2S: session-to-session accuracy; hea.: healthy; pat.: patients; dim. redu.: dimension reduction; spec.: specific; L/R/F: left/right/front; FS: feature selection; mdl.: model; sign.: significant; tt: total.

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007

Table 4
List of abbreviations.

Acronym	Full name	Acronym	Full name
CPCA	Class-wise PCA	PSD	Power spectral density
IDA	Independent discriminant analysis	CCR	Cross correlation
LAD	Laplacian derivation	ACC	Anterior cingulate cortex
LLSF	Laplacian spatial filter	OSF	Optimized spatial filter
ERSP	Event-related spectral perturbation	SCI	Spinal cord injury
BHDT	Bhattacharya distance	TPR	True positive rate
AME	Attempted ME	AIDA	Approximate IDA

line. A paired sample *t*-test using an alpha level of 0.05 for all statistical tests showed that the null hypothesis, i.e., the accuracies of two methods were of the same mean, was rejected for RMDR-JCFS ($M = 76.67$, $SD = 12.15$) versus: CSP ($M = 67.58$, $SD = 15.42$), $t(42) = 2.17$, $p = 0.0036$; FBPow ($M = 62.52$, $SD = 13.48$), $t(42) = 3.66$, $p = 0.0007$. However, the hypothesis was not rejected for RMDR-JCFS versus: FBCSP ($M = 71.64$, $SD = 15.04$), $t(42) = 1.22$; SWDCSP ($M = 69.64$, $SD = 14.35$), $t(42) = 1.75$; MD ($M = 74.79$, $SD = 12.40$), $t(42) = 0.51$; MDMR ($M = 72.78$, $SD = 11.65$), $t(42) = 1.08$.

To further demonstrate the effectiveness of our proposed method, i.e., RMDR-JCFS, its performance was compared with that of other channel selection methods, which were: common spatial pattern-based (cCSP) (Wang et al., 2005), fisher criterion-based (cFC) (Lal et al., 2004), mutual information-based (cMI) (Lan et al., 2005), zero-norm optimization-based (cZNO) (Lal et al., 2004) and recursive feature elimination-based (cRFE) (Lal et al., 2004). 8 channels were selected for other methods by considering performance of the algorithm with the comparison results shown in Table 2. It can be observed from the table, by jointly considering the frequency bands and channels, our proposed method performed significantly (i.e., using an alpha level of 0.05 for all statistical tests) better than most of the other existing methods by selecting the same number of channels. Paired sampled *t*-test showed that the accuracy of proposed method ($M = 76.67$, $SD = 12.15$), was significantly better than that of other channel selection methods: cCSP ($M = 65.35$, $SD = 14.73$), $t(42) = 2.78$, $p = 0.0081$; cMI ($M = 65.72$, $SD = 14.64$), $t(42) = 2.70$, $p = 0.01$; cFC ($M = 65.42$, $SD = 14.44$), $t(42) = 2.70$, $p = 0.0078$; cZNO ($M = 63.49$, $SD = 14.55$), $t(42) = 3.26$, $p = 0.0022$; and cRFE ($M = 65.70$, $SD = 13.95$), $t(42) = 2.78$, $p = 0.0081$. The subject-wise superior performance of proposed method can be noticed for each session, i.e., most best performances among all the methods were achieved by our proposed method. These results further validated the effectiveness of our proposed method.

4. Discussion and conclusions

Selection of the features was governed by the optimization/regularization presented in Eqs. (15) and (16) and schematically illustrated in Fig. 6, which consisted of three parts, i.e., the dependency of features with class labels ($\tilde{\mathcal{I}}_{wy}$), the redundancy between to-be-selected feature with those already selected ones ($\tilde{\mathcal{I}}_{ww}$) and the separability of two classes ($\tilde{\mathcal{I}}_{ss}$). These three items were jointly optimized/regularized. It was worth noting that the selected features could be different depending on how the first feature was selected. One way was to select it by maximizing $\tilde{\mathcal{I}}_{wy}$, which was used in our method. Only one feature would be selected if two features, e.g., A and B, carried similar $\tilde{\mathcal{I}}_{wy}$, $\tilde{\mathcal{I}}_{ww}$ and $\tilde{\mathcal{I}}_{ss}$. The prediction model built by including the selected feature, e.g., A, was expected to classify the discarded feature, e.g., B, correctly since they carried similar information. Actually, the dependency of the discarded feature, e.g., B, with those selected features should be increased with the inclusion of the selected feature A, hence, feature B may not be a good candidate any more. It was noted that the selection of the two features was in sequential processing order if they carried similar information.

The overall distributions of the selected channels and frequency bands for all the selected features across subjects were shown in Fig. 7. Observed from the figure, most selected channels were related to the activated areas in voluntary movements (VM) or MI of walking/foot dorsiflexion (Pfurtscheller and da Silva, 1999; Morash et al., 2008; Bakker et al., 2008; Pfurtscheller and Solis-Escalante, 2009; Muller-Putz et al., 2010; Do et al., 2011, 2013; Gwin et al., 2011; Cheron et al., 2012; Wang et al., 2012; Velu and de Sa, 2013; Castermans et al., 2014). The selection of 'F3', 'FC3' and 'FC4' agreed with the most activated areas such as prefrontal cortex, SMA and the leg and arm sensorimotor representation areas in alternating epochs of KMI and idling (Do et al., 2013). The premotor cortex was activated with significant changes from baseline to the phases of gait cycle in lower gamma band during active and passive robot-assisted walking (Wagner et al., 2012), which was also contributed to gait movements (Bakker et al., 2008). The selected 'FCz' and 'Cz' at the mid-central areas overlaying SMA were also involved in beta rebound in voluntary foot movement and imagery (Pfurtscheller and Solis-Escalante, 2009; Solis-Escalante et al., 2012). These were the most prominent areas in differentiating MI-Walking from idling at the *mu* and *low beta* bands (King et al., 2013). The involvement of 'Pz' in MI-Walking was supported by the significant differences in the foot area of sensory cortex (e.g., 'Pz' and 'Cz') between active and passive walking, and between active walking and resting in the *mu*, *beta* and *gamma* bands (Wagner et al., 2012). 'P3' and 'P4' were selected may be due to the visuo-motor integration and bimanual coordination involved during MI-Walking (Gwin et al., 2011), as supported by the critical dependence on the posterior parietal and motor cortex networks during flat and precision walking. The selection of 'C3' and 'C4' was consistent with the activated areas during voluntary limb movements or imagery at *mu* or *beta* band (Pfurtscheller and da Silva, 1999). While the selection of 'CP3' and 'CP4' agreed with the findings that the EEG powers in *mu* band at the lateral central ('C3' and 'C4') and lateral central-parietal ('CP3' and 'CP4') were the most informative in differentiating MI-Walking from idling (King et al., 2013). Selection of *mu*, *alpha*, and *beta* bands was supported by above discussions and the active frequency bands for ERD/ERS in VM and MI (Pfurtscheller and da Silva, 1999), and ME/MI of walking or foot dorsiflexion (Pfurtscheller and Solis-Escalante, 2009; Muller-Putz et al., 2010; Do et al., 2011, 2013; Gwin et al., 2011; Cheron et al., 2012; Wang et al., 2012; Velu and de Sa, 2013).

The frequent selection of 'Oz' may be due to the following reasons. 'Oz' may be shifted to 'Pz' considering the coarseness of the electrodes and elasticity of the cap, whereas 'Pz' was frequently activated during ME/MI-Walking or foot dorsiflexion (Wang et al., 2012; Wagner et al., 2012; Do et al., 2013; Velu and de Sa, 2013). The parieto-occipital and temporal-occipital visual association areas sub-served visual imagery (Roland and Gulyas, 1994) may be involved in MI-Walking attributed to the colorful background involved during MI-Walking/idle (see Fig. 1). Visualization of the walking movements while closing the eyes may activate the primary visual cortex (Kosslyn et al., 1995). Selection of 'Oz' can be

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007

explained by the association of the MI-Walking with the activation in the occipital visual areas (Jahn et al., 2004). An fMRI evidence showed that the extrastriate body area (EBA) were strongly modulated by arm and foot movements to a visual target stimulus and imagery activated some parts of EBA (Astafiev et al., 2004). Activation of 'Oz' may also attributed to the neural activity related to the activation of neurons coding for eye position relative to the orbits during eye movements (Law et al., 1998), i.e., the ocular artifacts.

To summarize, in this paper, we investigated the problem of detecting MI-Walking from background idle state for lower-limb rehabilitation of stroke patients. The power features filtered at multiple frequency bands were employed as features. The channels and frequency bands that contributed most to MI-Walking were jointly selected by optimizing the objective function formulated by: the dependency between features and class labels, the redundancy between the to-be-selected feature with those already selected ones, and the separation between two classes. Experimental results based on 11 healthy subjects yielded an averaged accuracy of 76.67%, which was 9.08%, 5.03%, 7.03%, 14.15% and 3.88% higher than that obtained by CSP, FBCSP, SWDCSP, FBPow and MDMR methods, respectively. Statistical tests with 95% confidence showed that only one out of eleven subjects performed at chance level. The effectiveness of the proposed method was further demonstrated by its superior performance compared with other channel selection methods. The averaged accuracy obtained by our proposed method was 9.12%, 8.75%, 9.05%, 10.98% and 8.77% significantly higher than that obtained by CSP, mutual information, fisher criterion, zero-norm optimization and recursive channel elimination-based methods, respectively. An averaged best session-to-session accuracy of 70.14% was obtained by proposed method.

Comparisons of our proposed method with relevant methods for the detection of motor imagery/movements execution of walking and foot dorsiflexion was presented in Table 3. A list of acronyms used in the table was shown in Table 4. The comparison revealed that detecting ME of walking or foot dorsiflexion (Morash et al., 2008; Muller-Putz et al., 2010; Do et al., 2011; Gwin et al., 2011; Cheron et al., 2012; Velu and de Sa, 2013) was relatively easier than that of MI (Morash et al., 2008; Pfurtscheller and Solis-Escalante, 2009; Muller-Putz et al., 2010; Niazi et al., 2011; Wang et al., 2012; Do et al., 2013), as evidenced by the differences in detection accuracies between ME and MI. The results demonstrated that the most frequently selected frequency bands relevant to MI-Walking, foot dorsiflexion were *alpha*, *mu* and *beta* bands (Do et al., 2011, 2013; Gwin et al., 2011; Cheron et al., 2012; Wang et al., 2012; Velu and de Sa, 2013; Castermans et al., 2014). While the most frequently activated or selected brain areas were located at the SMA, mid-central motor cortex of foot representation area, somatosensory motor cortex, primary motor cortex and posterior parietal sensorimotor cortex (Morash et al., 2008; Pfurtscheller and Solis-Escalante, 2009; Muller-Putz et al., 2010; Do et al., 2011, 2013; Gwin et al., 2011; Cheron et al., 2012; Wang et al., 2012; Velu and de Sa, 2013; Castermans et al., 2014).

The use of scalp EEG-based BCI beyond research lab for ambulatory context was rare, despite the need for cognitive assessments of human operators under real-world environment (Kerick et al., 2009). This would hamper the integration of BCIs with applications such as video gaming and virtual reality (Lotte et al., 2009; Kerick et al., 2009; Castermans et al., 2011, 2014; Duvinage et al., 2013). The event-related potentials examined under different motion conditions revealed that the percentage of acceptable trials and power spectrum of the signals were strongly affected by artifacts induced under the more dynamic conditions, e.g., walking and jogging (Kerick et al., 2009). Possible reasons on why limited studies existed for ambulatory conditions are as follows. EEG signals would be smeared by motions of body movements in ambulatory settings such as walking, swinging and jogging. The electromagnetic interference from lights and other potentials produced by eye blinks/movements, swallowing, and heart beat would pollute EEG signals. The triboelectric noise generated by movements, friction and flexion of cables was always exhibited in EEG signals under ambulatory conditions (Castermans et al., 2011). In addition, movements of electrodes produced by head movements and shocks modified the magnetic and capacitive coupling of the scalp with electrode leads (Kerick et al., 2009; Castermans et al., 2011). Advanced artifacts removal techniques are needed to separate the non-brain-relevant artifacts such as rhythmic activation corresponding to stride frequencies in walking or jogging from brain-relevant signals, which thus would improve the reliability of the signals (Kerick et al., 2009; Lotte et al., 2009). Deep understanding on how the motion-induced artifacts affect the interpretation of the observed signals relative to the underlying cognitive processing are required (Kerick et al., 2009).

References

Ang KK, Chin ZY, Wang C, Guan C, Zhang H. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. Front Neurosci 2012];6:1–9.
Ang KK, Guan C, Chua KSG, Ang BT, Kuah CK, Wang C, et al. A large clinical study on the ability of stroke patients to use EEG-based motor imagery brain–computer interface. Clin EEG Neurosci 2011];42:253–8.
Arvaneh M, Guan C, Ang KK, Quek C. Optimizing the channel selection and classification accuracy in EEG-based BCI. IEEE Trans Biomed Eng 2011];58:1865–73.
Astafiev SV, Stanley CM, Shulman GL, Corbetta M. Extrastriate body area in human occipital cortex responds to the performance of motor actions. Nat Neurosci 2004];7:542–8.
Bakker M, Overeem S, Snijders AH, Borm G, van Elswijk G, Toni I, et al. Motor imagery of foot dorsiflexion and gait: effects on corticospinal excitability. Clin Neurophysiol 2008];119:2519–27.
Belda-Lois J-M, Horno SMd, Bermejo-Bosch I, Moreno JC, Pons JL, Farina D, et al. Rehabilitation of gait after stroke: a review towards a top-down approach. J Neuroeng Rehabil 2011];8:1–19.
Blankertz B, Tomioka R, Lemm S, Kawanabe M, Muller KR. Optimizing spatial filters for robust eeg single-trial analysis. IEEE Signal Proc Mag 2008];25:41–56.
Castermans T, Duvinage M, Cheron G, Dutoit T. Towards effective non-invasive brain–computer interfaces dedicated to gait rehabilitation systems. Brain Sci 2014];4:1–48.
Castermans T, Duvinage M, Petieau M, Hoellinger T, Saedeleer C, Seetharaman K, et al. Optimizing the performances of a p300-based brain–computer interface in ambulatory conditions. IEEE J Emerg Sel Top Circ Syst 2011];1:566–77.
Chen ZP, Jiang JH, Li Y, Liang YZ, Yu RQ. Low-level segmentation of aerial images with fuzzy clustering. Chemom Intell Lab Syst 1999];45:295–302.
Cheron G, Duvinage M, Saedeleer CD, Castermans T, Bengoetxea A, Petieau M, et al. From spinal central pattern generators to cortical network: integrated BCI for walking rehabilitation. Neural Plast 2012], 375148.
Cui X, Jeter CB, Yang D, Montague PR, Eagleman DM. Vividness of mental imagery: individual variability can be measured objectively. Vis Res 2007];47:474–8.
Diaz I, Gil JJ, Sanchez E. Lower-limb robotic rehabilitation: literature review and challenges. J Robot 2011];2011:1–11.
Dickstein R, Deutsch JE. Motor imagery in physical therapist practice. Phys Ther 2007];87:942–53.
Dickstein R, Dunsky A, Marcovitz E. Motor imagery for gait rehabilitation in post-stroke hemiparesis. Phys Ther 2004];84:1167–77.
Do AH, Wang PT, King CE, Abiri A, Nenadic Z. Brain–computer interface controlled functional electrical stimulation system for ankle movement. J Neuroeng Rehabil 2011];8:1–14.
Do AH, Wang PT, King CE, Chun SN, Nenadic Z. Brain–computer interface controlled robotic gait orthosis. J Neuroeng Rehabil 2013];10:111.
Dunsky A, Dickstein R, Marcovitz E, Levy S. Home-based motor imagery training for gait rehabilitation of people with chronic poststroke hemiparesis. Arch Phys Med Rehabil 2008];89:1580–8.
Duvinage M, Castermans T, Petieau M, Hoellinger T, Seetharaman K, Cheron G, et al. A preliminary fundamental study of ambulatory SSVEP. In: Proc of the TOBI Workshop IV; 2013].
Farquhar J, Hill NJ, Lal TN, Schölkopf B. Regularised CSP for sensor selection in BCI. In: Proc of the 3rd Int Brain–Computer Interface Workshop and Training Course; 2006]. p. 14–5.
Gwin JT, Gramann K, Makeig S, Ferris DP. Electrocortical activity is coupled to gait cycle phase during treadmill walking. Neuroimage 2011];54:1289–96.
Jahn K, Deutschlander A, Stephan T, Strupp M, Wiesmann M, Brandt T. Brain activation patterns during imagined stance and locomotion in functional magnetic resonance imaging. Neuroimage 2004];22:1722–31.
Kerick SE, Oie KS, McDowell K. Assessment of EEG signal quality in motion environments. Army Res Lab Rep ARL-TR-4866; 2009].
Khushaba RN, Kodagoda S, Lal S, Dissanayake G. Driver drowsiness classification using fuzzy wavelet-packet-based feature extraction algorithm. IEEE Trans Biomed Eng 2011];58:121–31.

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007

King CE, Wang PT, Chui LA, Do AH, Nenadic Z. Operation of a brain–computer interface walking simulator for individuals with spinal cord injury. J Neuroeng Rehabil 2013];10:77.
Kosslyn SM, Thompson WL, Kim IJ, Alpert NM. Topographical representations of mental images in primary visual cortex. Nature 1995];378:496–8.
Lal TN, Schroder M, Hinterberger T, Weston J, Bogdan M, Birbaumer N, et al. Support vector channel selection in BCI. IEEE Trans Biomed Eng 2004];51:1003–10.
Lan T, Erdogmus D, Adami A, Pavel M, Mathan S. Salient EEG channel selection in brain–computer interfaces by mutual information maximization. Conf Proc IEEE Eng Med Biol Soc 2005];7:7064–7.
Law I, Svarer C, Rostrup E, Paulson OB. Parieto-occipital cortex activation during self-generated eye movements in the dark. Brain 1998];121:2189–200.
Lee HM, Chen CM, Chen JM, Jou YL. Low-level segmentation of aerial images with fuzzy clustering. IEEE Trans Syst Man Cybern 1986];16:589–98.
Lotte F, Fujisawa J, Touyama H, Ito R, Hirose M, Lecuyer A. Towards ambulatory brain–computer interfaces: a pilot study with p300 signals. In: Int Conf on Advances in Computer Entertainment Technology (ACE 2009); 2009]. p. 336–9.
Lotte F, Guan C. Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. IEEE Trans Biomed Eng 2011];58:355–62.
Malouin F, Richards CL, Durand A, Doyon J. Clinical assessment of motor imagery after stroke. Neurorehabil Neural Repair 2008];22:330–40.
Malouin F, Richards CL, Jackson PL, Lafleur MF, Durand A, Doyon J. The kinesthetic and visual imagery questionnaire (KVIQ) for assessing motor imagery in persons with physical disabilities: a reliability and construct validity study. J Neurol Phys Ther 2007];31:20–9.
Morash V, Bai O, Furlani S, Lin P, Hallett M. Classifying EEG signals preceding right hand, left hand, tongue, and right foot movements and motor imageries. Clin Neurophysiol 2008];119:2570–8.
Muller-Putz GR, Kaiser V, Solis-Escalante T, Pfurtscheller G. Fast set-up asynchronous brain-switch based on detection of foot motor imagery in 1-channel EEG. Med Biol Eng Comput 2010];48:229–33.
Niazi IK, Jiang N, Tiberghien O, Nielsen JF, Dremstrup K, Farina D. Detection of movement intention from single-trial movement-related cortical potentials. J Neural Eng 2011];8:066009.
Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell 2005];27:1226–38.
Pfurtscheller G, da Silva L. Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol 1999];110:1842–57.
Pfurtscheller G, Solis-Escalante T. Could the beta rebound in the EEG be suitable to realize a “brain switch”? Clin Neurophysiol 2009];120:24–9.
Ramoser H, Muller-Gerking J, Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movements. IEEE Trans Rehabil Eng 2000];8:441–6.
Roland PE, Gulyas B. Visual imagery and visual representation. Trends Neurosci 1994];17:281–7.
Schroder M, Lal TN, Hinterberger T, Bogdan M, Hill NJ, Birbaumer N, et al. Robust EEG channel selection across subjects for brain–computer interfaces. EURASIP J Appl Signal Process 2005];19:3103–12.
Severens M, Nienhuis B, Desain P, Duysens J. Feasibility of measuring event related desynchronization with electroencephalography during walking. Conf Proc IEEE Eng Med Biol Soc 2012]:2764–7.
Solis-Escalante T, Muller-Putz G, Brunner C, Kaiser V, Pfurtscheller G. Analysis of sensorimotor rhythms for the implementation of a brain switch for healthy subjects. Biomed Signal Process Control 2010];5:15–20.
Solis-Escalante T, Muller-Putz G, Pfurtscheller G. Overt foot movement detection in one single Laplacian EEG derivation. J Neurosci Methods 2008];175:148–53.
Solis-Escalante T, Muller-Putz G, Pfurtscheller G, Neuper C. Cue-induced beta rebound during withholding of overt and covert foot movement. Clin Neurophysiol 2012];123:1182–90.
Sun G, Hu J, Wu G. A novel frequency band selection in motor imagery based brain–computer interface. In: Proc of the IEEE World Congress on Comp. Intel. (WCCI 2010); 2010]. p. 335–40.
Tam WK, Tong KY, Meng F, Gao S. A minimal set of electrodes for motor imagery BCI to control an assistive device in chronic stroke subjects: a multi-session study. IEEE Trans Neural Syst Rehabil Eng 2011];19:617–27.
Velu PD, de Sa VR. Single-trial classification of gait and point movement preparation from human EEG. Front Neurosci 2013];7:84.
Villiger M, Estevez N, Hepp-Reymond M-C, Kiper D, Kollias SS, Eng K, et al. Enhanced activation of motor execution networks using action observation combined with imagination of lower limb movements. PLoS ONE 2013];8:e72403.
Wagner J, Solis-Escalante T, Grieshofer P, Neuper C, Muller-Putz G, Scherer R. Level of participation in robotic-assisted treadmill walking modulates midline sensorimotor EEG rhythms in able-bodied subjects. Neuroimage 2012];63:1203–11.
Wang PT, King CE, Chui LA, Do AH, Nenadic Z. Self-paced brain–computer interface control of ambulation in a virtual reality environment. J Neural Eng 2012];9:056016.
Wang Y, Gao S, Gao X. Common spatial pattern method for channel selection in motor imagery based brain–computer interface. In: Conf Proc IEEE Eng Med Biol Soc; 2005]. p. 5392–5.
Yang H, Guan C, Wang CC, Ang KK. Maximum dependency and minimum redundancy-based channel selection for motor imagery of walking EEG signal detection. In: Proc of the IEEE ICASSP; 2013]. p. 1187–91.
Yong X, Ward RK, Birch GE. Sparse spatial filter optimization for EEG channel reduction in brain–computer interface. In: Proc of the IEEE ICASSP; 2008]. p. 417–20.

Please cite this article in press as: Yang H, et al. Detection of motor imagery of brisk walking from electroencephalogram. J Neurosci Methods (2014), http://dx.doi.org/10.1016/j.jneumeth.2014.05.007