MAXIMUM DEPENDENCY AND MINIMUM REDUNDANCY-BASED CHANNEL SELECTION FOR MOTOR IMAGERY OF WALKING EEG SIGNAL DETECTION

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

Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore
Email: {hjyang, ctguan, ccwang, kkang}@i2r.a-star.edu.sg

ABSTRACT

This paper proposes a novel method to detect motor imagery of walking for the rehabilitation of stroke patients using the laplacian derivatives (LAD) of power averaged across frequency bands as the feature. We propose to select the most correlated channels by jointly considering the mutual information between the LAD power features of the channels and the class labels, and the redundancy between the LAD power features of the channel with that of the selected channels. Experiments are conducted on the EEG data collected for 11 healthy subjects using proposed method and compared with existing methods. The results show that the proposed method yielded an average classification accuracy of 67.19% by selecting as few as 4 LAD channels. An improved result of 71.45% and 73.23% are achieved by selecting 10 and 22 LAD channels, respectively. Comparison results revealed significantly superior performance of our proposed method compared to that obtained using common spatial pattern and filter bank with power features. Most importantly, our proposed method achieves significant better accuracy for poor BCI performers compared to existing methods. Thus, the results demonstrated the potential of using the proposed method for detecting motor imagery of walking for the rehabilitation of stroke patients.

Index Terms— rehabilitation, motor imagery of walking, EEG signal, mutual information, minimum redundancy.

1. INTRODUCTION

Stroke is one of the leading causes of mortality and disability in industralized countries [1]. One third of surviving patients lost the idenpendent walking ability and walking in an asymmetric manner. Gait recovery is one of the major objectives in the rehabilitation of stroke patients. Therapies and techniques used for gait rehabilitation include: classical techniques such as neurophysiological and motor learning approaches, functional electrical stimulation (FES), assistive robotic devices, and non-invasive brain-computer interfaces (BCI) [2, 3]. The primary objective of using BCI is to help the paralyzed patients to communicate with others and to improve their quality of life by volitional control of the brain activity. Similar to the robotic-based rehabilitation [1], BCI-based training allows more intensive repetitive motion and delivers the therapy with reasonable cost. It also allows the quantitative assessment of the level of recovery, which, hence, makes it possible to perform the training at home [3]. Motor imagery can even be used for the patients with no residual motor function [1]. The use of BCI for rehabilitation is strengthened by the followings. Firstly, it is anticipated that the rhythmic foot or leg movements activates the primary motor cortex, while the movement preparation activates the frontal and associated areas [1, 3]. The activation works in the comparable or similar way as that of the motor execution. Further, the activity-dependent plasticity throughout center nervous system will influence the functional outcome of the patients. Such mental imagery or rehearsal also causes the reorganization of the functional networks in both healthy and stroke people [3].

Classification techniques are usually used to detect the event-related synchronization (ERS) and event-related desynchronization (ERD) EEG signals generated by the preparation, mentally imaging and execution of the limbs. Most existing works focus on the detection of upper limb movement or imagination. Existing work on detection of lower limb movment and imagination includes: detection of dorsiflexion of both feet (movement) [4], comparsion of the effects for motor imagery of foot dorsiflexion and gait using motor evoked potentials and apply transcranial magnetic stimulation (tDCS) over the primary motor cortex [5], using the post-movement beta rebound after brisk feet movement to set up a classifier to classify the motor imagery of feet [6].

Recently, a large linical trial on 54 stroke patients have shown that the majority of stroke patients could use EEG-based BCI [7]. However, whether or not EEG-based BCI is helpful in gait rehabilitation still requires further investigation. In this paper, we address the issue based on the EEG data collected for 11 healthy subjects on motor imagery of walking and idle. We hope to investigate the followings: 1) Is it possible to detect the motor imagery of brisk walking with relative to a relaxation state? 2) what are the most correlated channels to the motor imagery of brisk walking? Can we detect the motor imagery of walking with the use of a few most correlated laplacian derivatives (LAD) channels? To achieve these objectives, we propose a method to select the most correlated channels by jointly considering the mutual informaiton between LAD power features of the channels and class labels, and the redundancy between LAD power features of the channel with that of the selected channels.

2. OUR PROPOSED METHOD

2.1. Laplacian Derivatives of Power Feature Extraction

Four electrodes horizontally or vertically neighboring to the selected electrode are employed to generate the Laplacian derivatives (LAD) power features. For those electrodes situated at the boundary, two electrodes horizontally or vertically neighboring to the selected electrode are used. The band power of laplacian derivatives of the chosen electrodes are used as the features. Let's firstly denote the signal as: $S_e(m,n,k)$, where m=1,2,...,N_s, n=1,2,...,N_c and k=1,2,...,N_r represent indexes of the samples, channels and trials, respectively. The signal is firstly divided into N_f frequency bands, starting from 4Hz to 44Hz, with the bandwidth of each frequency being 4Hz, and the overlapping between two frequency bands is taken as 2Hz. The signal is then band-pass filtered by Chebyshev filter, the resultant filtered signal is denoted as S_f. The band power for the kth trial, nth channel at frequency band f_s is calculated by

$$P_w(k,n,f_s) = 10 * log_{10}(\sum_{m=1}^{N_s} S_f(k,n,m) * S_f(k,n,m)) \tag{1}$$

Let's now assume the location indexes of the current processing electrode and its four neighboring electrodes as: $n_l(i,j)$, and $n_l(i\text{-}1,j)$, $n_l(i\text{+}1,j)$, $n_l(i,j\text{-}1)$ and $n_l(i,j\text{+}1)$, respectively; where i and j represent the locations in the x and y directions. In this case, the laplacian derivates of the power of the current channel (denoted by the location idex of $n_l(i,j)$) are given by

$$P_w^d(k,n_l(i,j),f_s) = P_w(k,n_l(i,j),f_s) - P_w^n(k,\hat{n}_l,f_s) \tag{2}$$

$$P_w^n(k,\hat{n}_l,f_s) = \frac{1}{4}\sum_{\substack{b=\{i-1,i+1\};\\ v=\{j-1,j+1\}}} P_w(k,n_l(b,v),f_s) \tag{3}$$

$$\bar{P}_w^d(k,n_l(i,j)) = \frac{1}{N_f}\sum_{f_s=1}^{N_f} P_w^d(k,n_l(i,j),f_s) \tag{4}$$

The final LAD power features ($\bar{P}_w^d(k,n_l(i,j))$) are the averaged power features across frequency bands.

2.2. Maximum dependency with minimum Redundancy (MD-MR)-based LAD Channel Selection

It is noticed that the most informative channels for generating the event-related de-synchronization (ERD) and event-related synchronization (ERS) are different from subject to subject. Further, these informative channels vary when the subject performs different mental tasks. To find out what are the most correlated channels for the motor imagery of walking action, we propose to choose the subject-specific LAD channels (electrodes) based on maximum dependency with minimum redundancy (MD-MR) as inspired by the feature selection method proposed in [8]. The maximum dependency is evaluated by the mutual information between the laplacian derivatives of the power features of the channels (w, here w is the feature ($\bar{P}_w^d$) obtained in Eq. (4)) and that of the class labels (c), i.e., the mutual information $I(w,c)$ should be the maximum in order to have the maximum dependency, which is given by

$$I(w;c) = \iint p(w,c) log\frac{p(w,c)}{p(w)p(c)}\,dw\,dc \tag{5}$$

where w_i and c_i represent the LAD power features for the LAD channel and that of the class labels; $p(w,c)$ represents the multivariate density. For the maximum dependency (MD)-based channel selection, the top N_d LAD channels that maximize $I(w;c)$ are selected, which is given by

$$\hat{n} = \arg\max_n \sum_{k=1}^{N_r} I(w(k,n),c(k)) \tag{6}$$

where k and n are the indexes of the trials and LAD channels. In general, the LAD channels can be selected in an incremental way, e.g., each time the one that maximizes the mutual information will be selected and added to the selected channel set.

However, the information contained in the selected channels may be redundant by only satisfying the maximum dependency condition. Hence, we hereby propose to impose the minimum redundancy constraints on the LAD power features of current considering channel with those already selected ones, i.e., the mutual information between the LAD power features of considering channel with those already selected ones should be minimum. This is achieved by jointly optimizing the condition

$$\min_{\substack{w(j)\in\\ \{W-W_{m-1}\}}} [\lambda\frac{1}{m-1}\sum_{\substack{w(i)\in\\ W_{m-1}}} I(w(j);w(i)) - I(w(j);c)] \tag{7}$$

where $w(i)$ is the brevity of $w(k,i)$. Each time, the LAD power features of selected channel should be of the minimum dependency with those already selected ones, e.g., those in W_{m-1}; λ is the weighting factor to balance the mutual information between the LAD power features of pairs of channels and that between the LAD power features of channels with class labels, λ=0.5 is selected in the experiments. The detail steps for the channel selection based on maximum dependency and minimum redundancy (MD-MR) are as follows.

Step 1. Compute the mutual information between the LAD power features of the channels with that of the class labels using Eq. (5).

Step 2. Choose a predefined number of features (N_p) that are of the maximum mutual information with the class labels using Eq. (6). This step is necessary if the original set is too large. N_p=α*N_f, where N_f is the final number of LAD channels to be selected, e.g., α=1.5 can be chosen.

Step 3. Compute the dependency between the LAD power features of each unselected channel with that of those already selected using Eq. (6).

Step 4. Choose the channels by jointly considering the dependency between the features of the channels with the class labels and the redundancy between LAD power features of unselected channels with those already selected channels using Eq. (7).

Repeat steps 3-4 till the required number of channels are selected.

3. EXPERIMENTAL RESULTS

To investigate the effectiveness of our proposed method in detecting the motor imagery of walking, EEG data were collected from 11 healthy subjects. The subjects are instructed to perform motor imagey of walking, i.e., to imagine walking using the two legs with the focus of the rhythmic movements of the legs, joints and the feelings when the feet touch and push the ground. Subjects gave written consent before participating the experiment, and none of the subjects has the history of neurological or orthopedic disorders. One trial consists of 16 seconds, with the preparation lasting for 2 seconds and shown as the changing of the traffic lights. The cue is then shown to the subjects as a human character walking or in stance position for 2 seconds. The subject is then asked to perform motor imagery of walking in his/her comfortable pace, or not to do anything by just looking at the screen. This is followed by a resting period of 6 seconds between any two trials. Two sessions of data are collected in two different days. Each session consists of two runs with each run consists of 40 trials of motor imagery of walking and 40 trials of idle. The subjects are of ages between 23-45, and among them, four are female and seven are male.

A total of 22 channels of laplacian derivatives (LAD channels) are symmetrically selected, which are: 'F3', 'F4', 'FC3', 'FC4', 'Fz', 'FCz', 'C3', 'C4', 'Cz', 'CPz', 'CP3', 'CP4', 'TP7', 'TP8', 'FT7', 'FT8', 'T7', 'T8', 'P3', 'P4', 'Pz' and 'Oz'. Among the 22 channels, 10 channels are further selected based on the proposed maximum dependancy with minimum redundancy (MD-MR)-based method. The optimal time segments for each subject are obtained using cross validation since the time points for actual peak responses are different from person to person. We compare the classification results of our proposed method of using 22 selected LAD channels with several existing methods, which include: filter bank common spatial pattern (FBCSP)-based EEG signal classification method [11], filter bank with power features (FBPF), common spatial pattern (CSP) method [12, 13] and sliding window discriminant CSP (SWD-CSP) method [14], with the 10×10 cross-validation results shown in Table 1. A paired t-test of the hypothesis that the difference of the two matched samples assumed to come from a normal distribution with mean zero has been rejected at significance of 95% for filter bank with power features (FBPF) and CSP, but not rejected for FBCSP and SWD-CSP. This indicates that our proposed method performs significantly better than filter bank with power features (FBPF) and CSP methods. Further, the averaged classification accuracy across subjects of our proposed method is 1.78% and 3.59% higher than that of FBCSP and SWD-CSP, respectively. Most importantly, our method has shown superior performance for those subjects of poor performance, e.g., subjects whose accuracies are below 70%, compared with other methods.

To further investigate how the performance is affected by selecting different numbers of channels based on our proposed maximum dependency with minimum redundancy (MD-MR)-based channel selection method, we further reduce the number of LAD channels used, e.g., using 4, 10 and 16 selected LAD channels, with the results shown in Fig. 1.

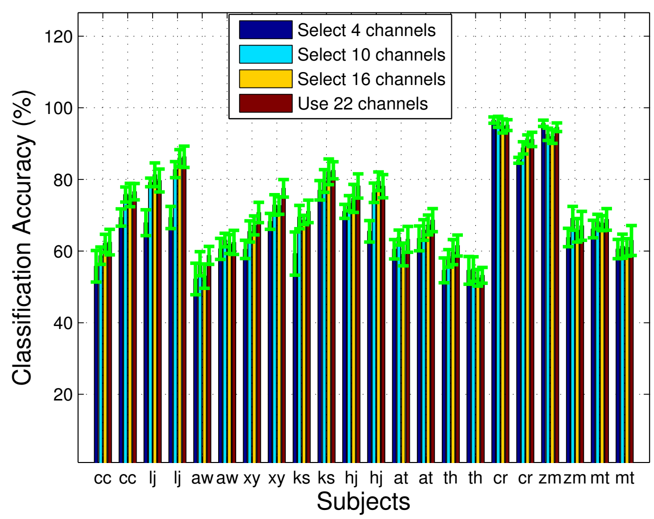

Fig. 1. Performance comparisons of selecting different numbers of LAD channels.

It can be observed from the figure that reducing the number of LAD channels only slightly degrades the performance compared with that achieved using all the 22 selected LAD channels. The average classification accuracies and variances obtained across subjects by selecting 4, 10, 16 LAD channels are: 67.19%±2.87, 71.45%±2.50, and 71.64%±2.67, which are slightly worse than that obtained using a total of 22 LAD channels, i.e., 73.23%±2.87. Hence, selecting the most informative channels will not degrade the performance much. However, the use of fewer channels will significantly reduce the setup time in the rehabilitation and hence, will make the system more practical to be used in clinical trials. The 95% confidence estimation of the accuracy for the re-

Table 1. Performance Comparisons of Proposed and Other Methods

Subj.	Sess.	Proposed (22 LAD chan.) $\mathcal{A}_c \pm \mathcal{V}_r$	Filter Bank with Power Fea. $\mathcal{A}_c \pm \mathcal{V}_r$	CSP $\mathcal{A}_c \pm \mathcal{V}_r$	SWD-CSP $\mathcal{A}_c \pm \mathcal{V}_r$	FBCSP $\mathcal{A}_c \pm \mathcal{V}_r$
cc	01	62.50±3.57	47.31±3.49	61.19±3.28	70.30±2.08	**73.06**±2.29
	02	**76.59**±2.28	48.75±5.85e-15	66.56±3.38	69.33±1.41	63.25±1.69
lj	01	79.70±3.18	68.56±2.32	75.56±2.73	77.61±1.59	**85.19**±2.19
	02	**86.32**±2.97	71.75±0.87	76.00±1.86	80.46±2.35	83.88±1.84
aw	01	58.79±2.50	56.13±2.46	**59.13**±2.28	53.32±5.82	53.50±3.46
	02	**62.41**±3.37	53.75±1.93	57.06±2.14	50.52±4.13	50.69±2.44
xy	01	**70.77**±2.83	50.69±3.29	55.88±2.60	62.53±1.41	58.38±1.29
	02	77.53±2.45	57.56±3.21	61.44±2.93	75.44±3.08	**78.25**±1.41
ks	01	71.08±3.13	62.00±1.95	51.94±3.25	70.53±2.92	**77.19**±2.52
	02	82.55±2.37	54.94±1.99	81.19±2.12	82.50±1.10	**85.81**±0.93
hj	01	78.17±3.37	77.50±0.78	85.25±1.72	**87.41**±1.49	87.38±1.90
	02	78.12±3.23	76.06±0.98	80.63±1.88	80.18±3.18	**85.31**±2.23
at	01	**63.28**±3.62	57.50±2.17	56.00±3.49	55.85±4.29	56.00±2.90
	02	68.64±3.22	53.85±3.22	60.80±1.97	57.29±1.16	**69.44**±2.56
cr	01	95.17±1.51	82.60±2.20	**96.81**±1.36	92.95±1.46	95.00±0.83
	02	91.17±1.98	89.81±0.59	**95.25**±0.91	91.29±1.35	90.56±1.04
zm	01	**94.57**±1.19	91.31±0.46	92.44±1.28	89.50±2.45	92.44±0.75
	02	**67.05**±4.02	64.69±1.11	64.69±2.97	66.34±2.87	65.13±1.76
mt	01	**68.84**±3.02	58.36±0.87	63.33±1.84	65.89±2.40	65.25±1.38
	02	**62.95**±4.16	49.34±1.71	52.06±2.30	48.74±4.57	48.94±4.65
th	01	**61.53**±2.94	53.63±1.69	47.88±2.98	50.67±1.08	56.00±1.87
	02	53.24±2.21	49.37±0.85	45.75±3.14	53.41±2.88	**55.44**±2.87
A_{ac}	*	73.23±2.87	62.52±1.73	67.58±2.38	69.64±2.50	71.45±2.04

A_c: Accuracy (%), $\mathcal{V}_r$: Variance. A_{ac}: Average Accuracy across subjects and sessions. Best performance is shown in bold.

spective action at chance level are: 41.88% and 57.50% using the inverse bionomial cumulative distribution. This indicates that subjects whose accuracy lie between 41.88% to 57.50% are deemed as performing at chance level. The results show that only subject 'th' at session 2 performed at chance level by using 22 LAD channels. Further analysis on the most correlated channels to the motor imagery of walking versus idle action shows that the most frequently activated locations across the subjects (e.g., in the sequence from the most to less frequently activated) are: 'P3', 'P4', 'Cz', 'CPz', 'T7', 'T8', 'Pz', 'Oz', 'C3', 'C4', 'FT7', 'FT8', 'F3', 'F4', 'CP3' and 'CP4', which are illustrated in Fig. 2. It

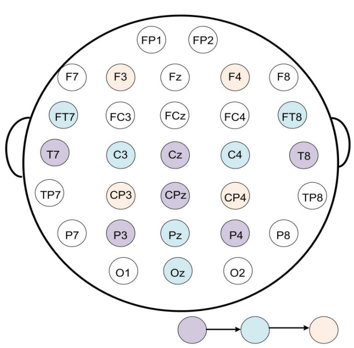

Fig. 2. An illustration of the most correlated locations for motor imagery of walking versus idle in a 10-20 EEG system.

can be further observed that the most correlated brain area to the motor imagery of walking action are concentrated in the medial primary sensorimotor cortices and the supplementary motor area, which is consistent with the evidence from positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) [10].

4. CONCLUSIONS

In this paper, we investigate the novel problem of detecting motor imagery of walking for stroke rehabilitation. The laplacian derivates (LAD) of power features for selected channels are used to eliminate the interference between electrodes. We propose to select the most correlated channels by jointly optimizing towards the maximum dependency between the features of LAD channels and class labels, and the minimum redundancy between the features of LAD channels with that of selected LAD channels. Experimental results demonstrate that by selecting as few as 4 LAD channels, we still achieve an acceptable classfication accuracy of 67.19%, which is slightly lower than 71.45% and 73.23% achieved by selecting 10 LAD channels and using 22 LAD channels. These results demonstrate that we can detect motor imagery of walking with fewer LAD channels. Comparisons with other methods demostrate significant better performance than filter bank with power features and CSP methods. Most importantly, our proposed method achieves superior performance for poor BCI performers compared with other methods. Statistical tests with 95% confidence show that only one out of eleven subjects performed at chance level. Our future work is to conduct clinical trials for the stroke patients.

5. REFERENCES

[1] I. Diaz, J. J. Gil, and E. Sanchez, "Lower-Limb Robotic Rehabilitation: Literature review and Challenges," *Journal of Robotics,* vol. 2011, pp. 1-11.

[2] J.-M. Belda-Lois, S. M.-d. Horno1, I. Bermejo-Bosch, J. C. Moreno, J. L Pons, D. Farina, M. Iosa, M. Molinari, F. Tamburella, A. Ramos, A. Caria, T. Solis-Escalante, C. Brunner and M. Rea, "Rehabilitation of gait after stroke: a review towards a top-down approach," *Journal of NeuroEngineering and Rehabilitation,* vol. 8, no. 66, pp. 1-19, 2011.

[3] A. Dunsky, R. Dickstein, E. Marcovitz and S. Levy, "Home-Based Motor Imagery Training for Gait Rehabilitation of People With Chronic PostStroke Hemiparesis," *Arch Phys Med Rehabil*, vol. 89, pp. 1580-1588, August 2008.

[4] T. Solis-Escalante, G. Muller-Putz and G. Pfurtscheller, "Overt foot movement detection in one single Laplacian EEG derivation," *Journal of Neuroscience Methods,* vol. 175, pp. 148-153, 2008.

[5] M. Bakker, S. Overeem, A. H. Snijders, G. Borm, G. van Elswijk, I. Toni and B. R. Bloem, "Motor Imagery of foot dorsiflexion and gait: Effects on corticospinal excitability," *Clinical Neurophysiology*, vol. 119, pp. 25199-2527, 2008.

[6] G. R. Muller-Putz, V. Kaiser, T. Solis-Escalane, G. Pfurtscheller, "Fast set-up asynchronous brain-switch based on detection of foot motor imagery in 1-channel EEG," *Med Biol Eng Comput*, vol. 11517, no. 9, pp. 572-577, January 2010.

[7] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang, C. W. K. Kuah, C. Wang, K. S. Phua, Z. Y. Chin and H. Zhang, "A Large Clinical Study on the Ability of Stroke Patients to Use EEG-Based Motor Imagery Brain-Computer Interface," *Clinical EEG and Neuroscience*, vol. 42, no. 4, pp. 253-258, October 2011.

[8] H. Peng, F. Long and C. Ding, "Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 27, no. 8, pp. 1226-1238, August 2005.

[9] S. d. Vries, M. Tepper, B. Otten and T. Mulder, "Recovery of Motor Imagery Ability in Stroke Patients, " *Rehabilitation Research and Practice*, vol. 2011, pp. 1-9.

[10] F. Malouin, C. L. Richards, "Mental Practice for Relearning Locomotor Skills," *Physical Therapy*, no. 90, pp. 240-251, Dec. 2009.

[11] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan and H. Zhang, "Filter Bank Common Spatial Pattern algorithm on BCI Competition IV Datasets 2a and 2b,", *Frontiers in Neuroscience,* vol. 6, pp. 1-9, March 2012.

[12] H. Ramoser, J. Muller-Gerking and G. Pfurtscheller, "Optimal Spatial Filtering of Single Trial EEG during Imagined Hand Movements," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 4, pp. 441-446, 2000.

[13] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Muller, "Optimizing Spatial Filters for Robust EEG Single-Trial Analysis," *IEEE Signal Proc. Magazine*, vol. 25, no. 1, pp. 41-56, 2008.

[14] G. Sun, J. Hu and G. Wu, "A Novel Frequency Band Selection in Motor Imagery Based Brain Computer Interface," *IEEE World Congress on Comp. Intel. (WCCI 2010)*, pp. 335-340, July 2010, Spain.