34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Online Semi-supervised Learning with KL Distance Weighting for Motor Imagery-based BCI

Atieh Bamdadian*, Cuntai Guan, Kai Keng Ang, and Jianxin Xu.

Abstract— **Studies had shown that Motor Imagery-based Brain Computer Interface (MI-based BCI) system can be used as a therapeutic tool such as for stroke rehabilitation, but had shown that not all subjects could perform MI well. Studies had also shown that MI and passive movement (PM) could similarly activate the motor system. Although the idea of calibrating MI-based BCI system from PM data is promising, there is an inherent difference between features extracted from MI and PM. Therefore, there is a need for online learning to alleviate the difference and improve the performance. Hence, in this study we propose an online batch mode semi-supervised learning with KL distance weighting to update the model trained from the calibration session by using unlabeled data from the online test session. In this study, the Filter Bank Common Spatial Pattern (FBCSP) algorithm is used to compute the most discriminative features of the EEG data in the calibration session and is updated iteratively on each band after a batch of online data is available for performing semi-supervised learning. The performance of the proposed method was compared with offline FBCSP, and results showed that the proposed method yielded slightly better results in comparison with offline FBCSP. The results also showed that the use of the model trained from PM for online session-to-session transfer compared to the use of the calibration model trained from MI yielded slightly better performance. The results suggest that using PM, due to its better performance and ease of recording is feasible and performance can be improved by using the proposed method to perform online semi-supervised learning while subjects perform MI.**

I. INTRODUCTION

Brain-computer interface (BCI) can be used as a communication and/or control system which enables the users to interact with external devices by their brain signals without any peripheral muscle activities [1]. BCI systems can also be served as a therapeutic tool to help people who suffer from motor impairments [2]. Motor impairment after stroke is the most important reason of permanent disability [3]. During the last decade several methods were developed to support stroke rehabilitation [3], [4], [5], [6], [7]: Active movement training (AMT), Electromyographic biofeedback, Robotics, and mental practice with Motor Imagery. Motor Imagery (MI) can be defined as a dynamic state during which the representation of a specific motor action is internally reactivated within working memory without any actual motor movement [3]. In other words, the motor system can be activated similarly during MI and actual movement [8]; therefore, MI would be useful approach for subjects with different levels of disabilities. However, using MI for rehabilitation has some problems: it is hard for the therapist to evaluate the performance of the MI performed by subject and also the subject himself has no feedback about his own performance. BCI system can help to overcome these problems [5]. Several previous studies used MI-based BCI for stroke rehabilitation [8], [9], [10]. MI based BCI can measure the brain activity and translate the imagination of the movements into commands. Moreover, the subject can be informed about his performance by means of a feedback and can try to enhance his performance relative to the received feedback.

Among different proposed methods for EEG preprocessing which is necessary for BCI systems, Common spatial pattern (CSP) algorithm [11] serves as an effective tool. CSP is used for discriminating the two classes of EEG data by maximizing the variance of one class while minimizing the variance of the other class. It has shown the performance of the spatial filters constructed by CSP algorithm depended on their operational frequency bands. Filter Bank Common Spatial Pattern (FBCSP) algorithm [12], [13] is one of the proposed methods which can automatically select the key temporal spatial discriminative EEG characteristics. Hence, due to the subject-specific frequency band selection by FBCSP higher performance is achieved in comparison with normal CSP. However, in session-to-session transfer and especially in online systems achieving a good performance is not possible by only selecting specific frequency band for each subject. The subject's brain signals may be changed from calibration session into test session. Therefore, there is a difference between calibration model and the online session model, which can be caused by several factors such as: subject's fatigue, getting involved in different tasks, or environmental interferences. In other words, the inherent non-stationarity behavior of the EEG signal can easily deteriorate the performance of the session-to-session transfer and also online EEG-based BCI.

It has shown in [14] that FBCSP can be applied for online adaptive and semi-supervised learning. In their study a Naïve Bayesian Parzen Window (NBPW) classifier is trained on the EEG data of the calibration session and used it to classify

*This work was supported by the Science and Engineering Research Council of A*STAR (Agency for Science, Technology and Research), Singapore

*A. Bamdadian is with Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, #21-01, Connexis, Singapore 138632, and Electrical and Computer Engineering Department, National University of Singapore, 10 Kent Ridge, Singapore 119260. (email: stuab@i2r.a-star.edu.sg).

C. Guan and K. K. Ang are with Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, #21-01, Connexis, Singapore 138632. (email:{ctguan, kkang}@i2r.a-star.edu.sg)

J. Xu is with Electrical and Computer Engineering Department, National University of Singapore, 10 Kent Ridge, Singapore 119260. (email:elecxujx@nus.edu.sg).

978-1-4577-1787-1/12/$26.00 ©2012 IEEE

the EEG data of the online test session. However, in their online and semi-supervised method only the NBPW classifier is updated. Here, in this study an online batch mode semi-supervised learning with KL distance weighting is applied to update FBCSP. The performance of our method is applied on the EEG data recorded from 12 healthy subjects. This paper is organized as follows: Section II describes the dataset and methodology of this study. Section III contains the experimental results, and finally section IV concludes the paper.

II. DATASET AND METHODS

A. Dataset

In this work, the EEG data recorded from the 12 healthy subjects are used. Two of the subjects were left handed and therefore perform motor imagery and passive movement by their left hand while the rest perform by their right hand. All the subjects were asked for ethics approval and informed consent. EEG signal were recorded by using the Nuamps EEG acquisition hardware (http://www.neuroscan.com) with unipolar Ag/AgCl electrodes channels, digitally sampled at 250 Hz with a resolution of 22 bits for voltage ranges of 130 mV. EEG recordings from all 27 channels are bandpass filtered from 0.05 to 40 Hz by the acquisition hardware. All subjects were asked to minimize any physical movement and eye blinking throughout the EEG recording process.

The EEG data from each subject were collected on two separate days. On the first day four non feedback sessions were recorded. The first two sessions collected EEG from a subject while performing motor imagery of the chosen hand and background rest condition. During these two sessions, the subjects were instructed to perform kinaesthetic motor imagery of their chosen hand right after a visual cues displayed on the computer screen in each trial. During the background rest condition, the subjects were instructed to perform mental counting. This instruction was given to define the background rest condition to the subject. The next two sessions collected EEG data from the subject while passive movement of the chosen hand was performed using the haptic knob robot [6] and background rest condition. During these two sessions, the subjects were supposed to be relax while the movement of the chosen hand was performed using the haptic knob robot [6]. During the background rest condition subjects performed mental counting similar to the first two sessions.

Each session lasted about for approximately 16 minutes that comprised of 40 trials of either motor imagery or passive movement, and 40 trials of background rest condition. Each trial comprised a preparatory segment of 2s, the presentation of the visual cue for 4s, and a rest segment of at least 6s. Each trial lasted approximately 12s, and a break period of at least 2 minutes was given after each session of EEG recording. To calibrate the subject-specific model from performing motor imagery (passive movement), the first (last) two sessions were used.

On the second day, four sessions of EEG data were collected with feedback from the subjects while performing

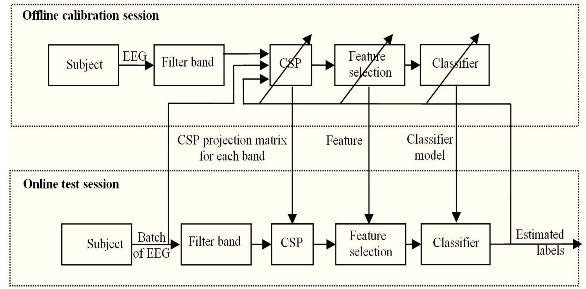

Fig. 1: Architecture of online batch mode semi-supervised learning method based on Filter Bank Common Spatial Pattern (FBCSP) algorithm.

motor imagery of the chosen hand and background rest condition. In the first two sessions of the second day the BCI system was calibrated by MI calibration session data, and in the next two sessions the BCI was calibrated by PM calibration session. Each session again lasted about for approximately 16 minutes that comprised of 40 trials of motor imagery and 40 trials of background rest condition.

B. Filter Bank Common Spatial Pattern (FBCSP)

In this study, we use Filter Bank Common Spatial Pattern (FBCSP) algorithm [12], [14] FBCSP can effectively select the subject specific frequency band for the CSP. It has four progressive stages of EEG measurements processing [13]:

- Multiple frequency band pass filtering: A total of 9 Chebyshev Type II band-pass filters are used, namely, 4-8 Hz, 8-12 Hz, ... , 36-40 Hz.
- Spatial filtering: The CSP algorithm is applied to spatially filter the signal. The CSP features are computed for each band-pass frequency range.
- Features selection: Four best features are selected among all 36 features by using the Mutual Information-based Best Individual Feature (MIBIF) algorithm.
- Classification: SVM classification algorithm is employed to model and classify the selected CSP features.

More details about the FBCSP algorithm can be found in [12], [13].

C. Batch mode updating of FBCSP

online learning for MI-based BCI system, the subject may instruct to perform specific motor imagery action; this means the labels of the online test session are known. However, in real BCI applications it is more preferable that the subject be free to perform either types of motor imagery action. In such later cases, the labels would be unknown, since the subject is not given any instructions to perform specific motor imagery action. Semi-supervised learning is a type of adaptive learning deals with these cases where both labeled and unlabeled data are available [15]. In this study, we have labeled passive movement data from offline calibration session to train the classifier; the newly recorded online test session data will be used in batches to update FBCSP iteratively, which means that in each iteration the training data set, CSP feature vectors of training data, and the labels of each batch of trials from the test data are updated. The online batch mode semi-supervised algorithm, illustrated in Fig. 1, is described in the following:

- Step 1: Use train data to train FBCSP and estimate the labels of the newly recorded batch of test data. Since it is online system these estimated labels shall be kept for accuracy estimation.
- Step 2: For kth iteration (k=1:K):
  Add the newly recorded batch of test data with estimated labels to the train data. Re-train FBCSP to derive the new features, and re-estimate the labels of the newly recorded test data.
- Step 3: Update the train model by adding the newly recorded test data with the estimated labels after Kth iteration.
- Step 4: Go to Step 1

According to our simulations and also similar to some other literature [16], the number of iterations were fixed to K=3. Due to the total number of test trials we planned to record and the speed of the online system we choose 10 trials in each batch.

Various semi-supervised methods had proposed previously [16], [17]. They used Expectation Maximization (EM) algorithm iteratively, and the train data is augmented with the predicted labels of the current evaluation data. In [14] they use some confidence level to avoid adding all trials, they just augment the train data by those trials which match the probabilistic model captured by NBPW. Also, in [18] they used adaptive feature extraction and assign the variability coefficient manually. Here in this study, we update the CSP of each band iteratively. This can be done by increasing the number of trials. Therefore, the covariance of each class in each band is updated as follows:

$$\Sigma_{b,(c)} = (1-\alpha)\Sigma^{tr}_{b,(c)} + \alpha\Sigma^{ts}_{b,(c)}, \quad (1)$$

where $\Sigma^{tr}_{b,(c)}$ represents the covariance matrix of class c in b^{th} band, and tr and ts denotes the covariance of train and test data, respectively. The weighting parameter α is defined $\alpha = \frac{n}{N+n}$, where N is the total number of trials of passive movement calibration data and n is the total number of recorded trials from test data till now. This idea is similar to updating CSP in Composite Common Spatial Pattern (CCSP) [19]. In CCSP the CSP is updated by using the data from other subjects. Here, the CSP on each band is updated using the newly recorded data from the test session.

In this paper, we propose online batch mode semi-supervised learning with KL distance weighting which has a structure explained above but assigns some extra weights to the incoming trials from the test data:

$$\Sigma_{b,(c)} = (1-\alpha)\Sigma^{tr}_{b,(c)} + \frac{1}{KL_{b,(c)}}\alpha\Sigma^{ts}_{b,(c)}, \quad (2)$$

$$\begin{aligned} KL_{b,(c)} =& 0.5\{\log(\frac{\det \Sigma^{tr}_{b,(c)}}{\det \Sigma^{ts}_{b,(c)}}) + \\ & trace((\Sigma^{tr}_{b,(c)})^{-1}\Sigma^{ts}_{b,(c)}) - D\}, \end{aligned} \quad (3)$$

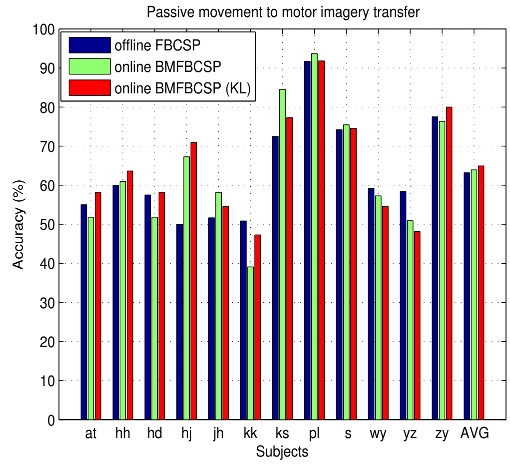

Fig. 2: The accuracies of motor imagery detection (with feedback) calibrated by passive movement using offline (offline FBCSP), online batch mode semi-supervised (online BMFBCSP), and online batch mode semi-supervised method with KL distance weighting (online BMFBCSP (KL)).

where $KL_{b,(c)}$ is a Kullback-Leibler (KL) [20] distance shows the difference between probability distribution of the train and incoming trials from test data up to now, D is the dimension of the covariance matrix and det represents determinant of a matrix.

III. RESULTS AND DISCUSSION

This section evaluates the performance of the proposed method so called online batch mode semi-supervised learning with KL distance weighting in comparison with offline method for MI-based BCI (with feedback) calibrated by passive movement data. The proposed method is applied on the data set with 12 healthy subjects described in section II-A.

Fig. 2 shows the accuracies of the PM to MI transfer in three different experiments: 1) offline FBCSP, 2) online batch mode semi-supervised method (BMFBCSP) in which the train model is updated iteratively by adding the new batch of recorded test data, and 3) online BMFBCSP with KL distance weighting which adds the weighted new batch of recorded test data to update the train model iteratively. As can be seen, the average accuracy over 12 subjects is increased for both BMFBCSP (63.94%) and BMFBCSP with KL distance weighting (65.91%) in comparison with offline FBCSP (63.19%). The results of using BMFBCSP with KL distance weighting also indicate that for 9 out of 12 subjects we get better accuracy in comparing to offline analysis. In offline analysis there is no usage of current test session. This suggests that adaptation in our problem is helpful to alleviate the difference between the calibration session and the test session. As an example subject hj who have a random performance in offline analysis, shows about 20% improvement by using online adaptation.

Among the 12 subjects, five subjects (i.e., hh, kk, ks, s, zy) had some prior experiments in operating MI-based BCI

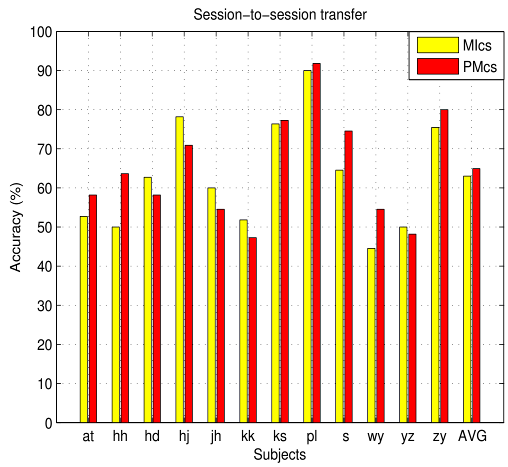

Fig. 3: The accuracies of motor imagery (with feedback) detection using passive movement (PMcs) or motor imagery (MIcs) for calibration using online batch mode semi-supervised method with KL weighting.

while the rest were BCI-naïve subjects. As expected, the average accuracies of the BCI-naïve subjects for all three experiments are around 6% less than the average accuracies of the subjects with prior BCI experiences. The 10x10 cross validation accuracies of detecting passive movement/ motor imagery from the background rest condition are calculated and was also previously reported in [21]. Those subjects with accuracies above 80 are considered as the best subjects (i.e., hj, jh, kk, pl, s, zy). The average accuracy using BMFBCSP with KL distance weighting (70.15%) over these best subjects increased about 5% in comparing to the offline FBCSP (65.97%).

For better evaluation of the proposed online semi-supervised method (BMFBCSP) with KL distance weighting, it is also applied for MI detection with feedback calibrated by MI calibration session and compared with the results calibrated by PM calibration session (Fig. 3). As shown, using the proposed method for online MI detection calibrated by PM has slightly better performance than calibrated by MI. Moreover, the performance of 7 out of 12 subjects was better when calibration is performed by PM in online MI detection. For the rest, the drops could be still because of the difference between MI and PM calibration sessions.

IV. CONCLUSION

Motor imagery-based BCI systems are applied in different applications. However there is a common challenge which affects their performances. It is not possible for all the subjects either disabled or healthy to perform motor imagery correctly. One of the feasible proposed methods to overcome this challenge is to calibrate the system with passive movement data. This study indicated that the MI-based BCI system calibrated with PM can be used for online MI detection with feedback. The results showed that on average the performance of the MI-based BCI system (with feedback) calibrated by PM data in online system is slightly better than the one calibrated with MI in both offline and online systems. The results may improve more by using some advanced adaptation methods to overcome the difference between calibration sessions of MI and PM. However, the results are promising enough to suggest applying the proposed method for online MI detection in therapeutic or non-therapeutic applications.

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," Clini. Neurophysiol., vol. 113, no. 6, pp. 767-791, 2002.

[2] J. N. Mak and J. R. Wolpaw, "Clinical Applications of Brain-Computer Interfaces: Current State and Future Prospects," IEEE Rev. Biomed.l Eng., vol. 2, pp. 187-199, 2009.

[3] N. Sharma, V. M. Pomeroy, and J.-C. Baron, "Motor Imagery: A Backdoor to the Motor System After Stroke?," Stroke, vol. 37, pp. 1941-1952, Jul. 2006.

[4] L. Kalra, "Stroke Rehabilitation 2009: Old Chestnuts and New Insights," Stroke, vol. 41, pp. e88-e90, Feb. 2010.

[5] J. J. Daly, R. Cheng, K. H. J. M. Rogers, and K. L. M. E. Dohring,"Development and Testing of Non-Invasive BCI + FES/Robot System For Use in Motor Re-Learning After Stroke," in Proc. 13th Annu. Conf. Int. Functional Elect. Stim. Soc., Sept. 2008, pp. 200-202.

[6] O. Lambercy, et al., "A Haptic Knob for Rehabilitation of Hand Function," IEEE Trans. Neural Syst. Rehabi. Eng.,vol. 15, pp. 356-366, 2007.

[7] D. Intiso, V. Santilli, MG. Grasso, R. Rossi, and I. Caruso, "Rehabilitation of walking with electromyographic biofeedback in foot-drop after stroke," Stroke, vol. 25, pp. 1189-1192, Jun. 1994.

[8] V. Kaiser, A. Kreilinger, G. R. Muller-Putz, and C. Neuper, "First steps towards a motor imagery based stroke BCI New strategy to set up a classifier," Frontiers in Neuroscience, vol. 5, pp. 1-10, Jul. 2011.

[9] K. K. Ang, et al., "Clinical study of neurorehabilitation in stroke using EEG-based motor imagery brain-computer interface with robotic feedback," in Proc. 32nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2010, pp. 5549-5552.

[10] K. K. Ang, et al., "A Large Clinical Study on the Ability of Stroke Patients to Use an EEG-Based Motor Imagery Brain-Computer Interface," Clin. EEG and Neuroscience, vol. 42, pp. 253-258, 2011.

[11] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Muller, "Optimizing Spatial filters for Robust EEG Single-Trial Analysis," IEEE Signal Process. Mag., vol. 25, pp. 41-56, 2008.

[12] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface," in Proc. IEEE Int. Joint Conf. Neural Netw., 2008, pp. 2390-2397.

[13] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Mutual information-based selection of optimal spatial-temporal patterns for single-trial EEG-based BCIs," Pattern Recog., vol. 45, pp. 2137-2144, 2012.

[14] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter Bank Common Spatial Pattern (FBCSP) algorithm using online adaptive and semi-supervised learning," in Proc. IEEE Int. Joint Conf. Neural Netw., 2011, pp. 392-396.

[15] I. Cohen, F. G. Cozman, N. Sebe, M. C. Cirelo, and T. S. Huang, "Semisupervised learning of classifiers: theory, algorithms, and their application to human-computer interaction," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, pp. 1553-1566, 2004.

[16] Y. Li and C. Guan, "An extended EM algorithm for joint feature extraction and classification in brain-computer interfaces," Neural Comput., vol. 18, pp. 2730-2761, Nov. 2006.

[17] Y. Li and C. Guan, "Joint feature re-extraction and classification using an iterative semi-supervised support vector machine algorithm," Mach. Learning vol. 71, pp. 33-53, Apr. 2008.

[18] S. Sun and C. Zhang, "Adaptive feature extraction for EEG signal classification," Med Bio Eng Comput, vol. 44, pp. 931-935, 2006.

[19] K. Hyohyeong, N. Yunjun, and C. Seungjin, "Composite Common Spatial Pattern for Subject-to-Subject Transfer," Signal Processing Letters, IEEE, vol. 16, pp. 683-686, 2009.

[20] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern classification: John Wiley and Sons Inc, 2001.

[21] K. K. Ang, et al., "Calibrating EEG-based motor imagery brain-computer interface from passive movement," in Proc. 33nd Annu. Conf. IEEE Eng. Med. Soc., 2011, pp. 4199-4202.