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

Omitting the Intra-session Calibration in EEG-based Brain Computer Interface Used for Stroke Rehabilitation

Mahnaz Arvaneh, Cuntai Guan, Kai Keng Ang and Chai Quek

Abstract— **Brain-computer interface (BCI) as a rehabilitation tool has been used in restoring motor functions in patients with moderate to sever stroke impairments. To achieve the best possible outcome in such an application, it is highly desirable to have a stable and accurate operation of BCI. However, since electroencephalogram (EEG) signals considerably vary between sessions of even the same user, typically a long calibration session is recorded at the beginning of each session. This process is time-consuming and inconvenient for stroke patients who undergo long-term BCI sessions with repeating same mental tasks. This paper investigates the possibility of omitting the intra-session calibration for BCI-based stroke rehabilitation when large data recorded from the same user are available. For this purpose, a large dataset of EEG signals from 11 stroke patients performing 12 BCI-based stroke rehabilitation sessions over one month is used. Our offline results suggest that after recording a number of stroke rehabilitation sessions, the patient does not require calibration any more. The experimental results show that combining 11 sessions, which each session comprises minimum 60 trials per class, yields a model that averagely outperforms the standard calibration model trained by the data recorded directly before the test session.**

I. INTRODUCTION

Brain-computer interface (BCI) provides a nonmuscular means of communication and controlling a device [1]. Through voluntary motor imaginations or movement intentions, brain activities can be decoded to controlling signals. Thus, BCI technology enables people with sever motor disabilities to use their brain signals for communication and control [2]. Furthermore, as a rehabilitation tool BCI has been effectively used in restoring motor functions in patients with moderate to sever stroke impairments [3]. In such a system, BCI could guide brain plasticity by demanding close attention to a motor task or by requiring the activation or deactivation of specific brain signals.

In majority of current BCI systems, the brain signals are measured by EEG, due to its low cost and high time resolution [4]. Since, the EEG patterns considerably vary between sessions even for a same subject, the motor imagery-based BCI typically requires to record labelled training data during a so called calibration phase at the begin of each session. The calibration phase takes around 20-30 minutes, and its acquired data are used to train a BCI model adapted to the subject's current EEG patterns. This time-consuming preparation step is especially inconvenient for patients who are long-term BCI users. Therefore, the question arises whether we can omit the time-consuming calibration phase. In the other words, is it possible to use the recorded data from the previous sessions of the same subject to train a stable BCI model that can be reliably used in follow-up sessions?

Krauledat et al. recently proposed an algorithm to skip the calibration phase of long-term BCI users by concatenating and clustering the historic spatial filters of the same user [5]. The previous findings in [5] were further confirmed in an online study published in [6]. More recently, Fazli et al. proposed a method to omit the calibration phase for long-term and novel subjects by an ensemble of historic sessions [7].

While stroke patients are potential audiences of BCI who repeatedly perform BCI sessions with same mental tasks, all the aforementioned studies focused only on healthy and expert BCI subjects. To the authors' knowledge, so far there are no studies investigating the stability and robustness of a stroke rehabilitation system without an intra-session calibration. fMRI and PET studies on stroke rehabilitation revealed dynamic changes in the activation patterns during recovery [8]-[10]. Therefore, the reliability of transferring information from prior BCI-based stroke rehabilitation sessions needs to be investigated.

This study aims to investigate whether combining different stroke rehabilitation sessions of a same subject allows us to reveal the "invariant" BCI features. For this purpose, a 12 sessions motor imagery-based BCI dataset collected from 11 stroke patients over one month [3], [11] is used. In this study, the Filter Bank Common Spatial Pattern (FBCSP) algorithm which is the basis of all wining algorithms in BCI competition IV is applied to train the subject-specific models [12], [13]. We show how many data from different sessions are required to train a good classifier which eliminates the necessity of going through an intra-session calibration for a new session. Moreover, the trained model by the historic data is tested against the standard model trained by the data recorded directly before the test session.

The remainder of this paper is organized as follows: Section II briefly describes the FBCSP algorithm. The applied dataset and the performed experiments are explained in Section III. Section IV presents the experimental results and finally Section V concludes the paper.

This work was supported by the Agency for Science, Technology and Research (A.STAR), and the Nanyang Technological University, Singapore.
*M. Arvaneh is with School of Computer Engineering, Nanyang Technological University, Singapore 639798, and Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), Singapore 138632. (email: stuma@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), Singapore 138632. (email: kkang, ctguan@i2r.a-star.edu.sg).
C. Quek is with School of Computer Engineering, Nanyang Technological University, Singapore 639798. (e-mail: ashcqek@ntu.edu.sg).

II. METHODOLOGY

A. Filter Bank Common Spatial Pattern

The common spatial pattern (CSP) algorithm has been shown to be effective in discriminating two classes of motor imagery tasks [14]. Despite its effectiveness and widespread use, the CSP performance greatly depends on its operational frequency band [14], [15]. The FBCSP algorithm addresses this issue by performing consecutively multi-band spectral filtering and common spatial filtering to extract and select most discriminative features.

Due to its superior performance against CSP, this study used the filter bank common spatial pattern (FBCSP) algorithm [12], [13]. The FBCSP methodology comprises the following steps:

Step 1) Multi-band spectral filtering: The first step uses a filter bank that decomposes the EEG data using nine equal bandwidth filters, namely 4-8, 8-12, ..., 36-40 Hz as proposed in [12], [13]. These frequency ranges cover most of the manually or heuristically selected settings used in the literature.

Step 2) common spatial filtering: In this step, the EEG data from each frequency band are spatially filtered using CSP filters [14]. Let $\mathbf{X_b} \in \mathbf{R}^{N_c \times S}$ denote a single-trial EEG data from the b^{th} band-pass filter, where N_c and S denote the number of channels and number of measurement samples respectively. A linear projection transforms $\mathbf{X_b}$ to spatially filtered $\mathbf{Z_b}$ as

$$\mathbf{Z_b} = \mathbf{w}_\mathbf{b}^* \mathbf{X_b}, \tag{1}$$

where each row of the transformation matrix $\mathbf{w}_\mathbf{b}^* \in \mathbf{R}^{2m \times N_c}$ indicates one of the $2m$ CSP filters.

The CSP transformation matrix corresponding to the b^{th} band-pass filter, $\mathbf{W}_b$, is generally computed by solving the eigenvalue decomposition problem:

$$\mathbf{C}_{b,1}\mathbf{W}_b = \left(\mathbf{C}_{b,1} + \mathbf{C}_{b,2}\right)\mathbf{W}_b\mathbf{D}, \tag{2}$$

where $\mathbf{C}_{b,1}$ and $\mathbf{C}_{b,2}$ are respectively the averaged covariance matrices of the band-passed EEG data of each class; $\mathbf{D}$ is the diagonal matrix that contains the eigenvalues of $(\mathbf{C}_{b,1} + \mathbf{C}_{b,2})^{-1}\mathbf{C}_{b,1}$. Usually, only the first and last m rows of $\mathbf{W}_b$ are used as the most discriminative filters to perform spatial filtering [14].

Step 3) Feature extraction: The spatio-spectrally filtered EEG data are used to determine the features associated to each band-pass frequency range. Based on the Ramoser formula [16], the features of the k^{th} trial of the EEG data from the b^{th} band-pass filter are given by

$$\mathbf{v}_{b,k} = \log(\text{diag}(\mathbf{Z}_{b,k}\mathbf{Z}_{b,k}^{\text{T}})/\text{trace}[\mathbf{Z}_{b,k}\mathbf{Z}_{b,k}^{\text{T}}]), \tag{3}$$

where $\mathbf{v}_{b,k} \in \mathbf{R}^{1\times 2m}$; diag(.) returns the diagonal elements of the square matrix; trace[.] returns the sum of the diagonal elements of the square matrix; and the superscript T denotes the transpose operator. Since we have nine frequency bands, the feature vector for the k^{th} trial is formed using

$$\mathbf{V}_k = [\mathbf{v}_{1,k}, \mathbf{v}_{2,k}, ..., \mathbf{v}_{9,k}]. \tag{4}$$

where $\mathbf{V}_k \in \mathbf{R}^{1\times 18m}$.

Step 4) Feature selection: The last step selects the most discriminative features of the feature vector $\mathbf{V}$. Various feature selection algorithms can be used in this step. Based on the results presented in [12] the Mutual Information-based Best Individual Feature is used to select four pairs of features.

III. EXPERIMENTS

A. Data Description

In this study, the EEG data from 11 hemiparetic stroke patients who underwent motor imagery-based BCI with robotic feedback neuro-rehabilitation were used (refer NCT00955838 in ClinincalTrials.gov) [3]. Each patient underwent 12 sessions of 1-hour neuro-rehabilitation on the stroke-affected hand for one month. Table I provides more clinical information about the 11 stroke patients.

TABLE I
DEMOGRAPHIC AND CLINICAL INFORMATION FOR N=11 STROKE SUBJECTS WHO PARTICIPATED IN THIS STUDY

Gender M/F (%)	Handedness R/L (%)	Stroke Type I/H (%)	Stroke Side R/L (%)	Stroke Nature C/S (%)	Mean age (Range)	CVA to therapy days (Range)	Week 0 FMA (Range)
9M	10R	5I	6R	3C	47.5±13.5	383±291	26.3±10.3
(81.8)	(90.9)	(45.5)	(54.5)	(27.3)	(23-61)	(71-831)	(17-47)

M indicates Male; F, Female; R, Right; L, Left; N, None; I, Infarction; H, Haemorrhagic; C, Cortical; S, Subcortical; CVA, Cerebrovascular accident; FMA, Fugl- Meyer Assessment; N, Number of patients.

In the rehabilitation phase, the patient's impaired hand was strapped to the MIT-Manus robot. Thus, when a successful motor imagery task was completed, the MIT-Manus robot passively moved the stroke-affected limb. The EEG signals were acquired using Nuamps acquisition hardware (http://www.neuroscan.com) with 27 unipolar Ag/AgCl electrodes channels, digitally sampled at 250 Hz with a resolution of 22 bits for voltage ranges of 130 mV. The recorded EEG signals were band-pass filtered from 0.05 to 40 Hz by the acquisition hardware.

Fig. 1 shows the timing for the rehabilitation session. In each trial the patient was first prepared with a visual cue for 2 s, then a go cue instructed the patient to perform motor imagery of the impaired hand. If the voluntary motor intent was detected within the 4 s action period, the MIT-Manus robot assisted the patient to move the impaired hand towards the goal. Finally the patient was asked to rest for 6 s. Each patient underwent 12 neuro-rehabilitation sessions, 3 sessions per week. There was a total of 160 repeats in each session (1 repeat means a complete run from preparation cue to the rest stage). There was a dedicated calibration phase before the rehabilitation phase to train the online classifier.

B. EEG signal processing

In our study, the classification problem involved distinguishing between the motor imagery stage and the rest stage. Therefore, each session comprised 160 motor imagery actions of the affected hand and 160 rest conditions. The rest

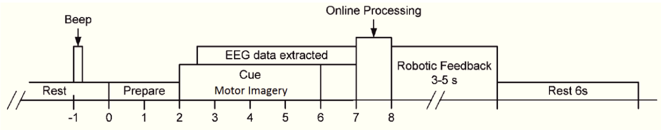

Fig. 1. Timing of each repeat in the performed neuro-rehabilitation phase with on-line robotic feedback.

class is a "no-command" state that the patients were allowed to do almost any other mental tasks than the impaired hand motor imagery. There were a total of 12 sessions of BCI data recorded on different days for each patient.

An overview of the performed experiments are shown in Fig. 2 and Fig. 3. These experiments aimed to evaluate the performance of a model trained using the other stroke rehabilitation sessions of the same subject, and compare it against the standard model trained on a part of the new session.

In the first experiment, as shown in Fig. 2, we trained models using different portions of 11 sessions (m% which m varies from 2.5 to 100), and tested on the second half of the test session. Since we assumed that the collected sessions of a subject are independent, in each run a session (from the 12 available sessions of a subject) was considered as the test session and the rest of 11 sessions were used to train a model. The model was trained using different numbers of trials selected equally and randomly from each session. This experiment investigated whether combining different stroke rehabilitation sessions of a same subject recorded over one month allows us to reveal stable BCI models. Moreover, this experiment considered that how many trials from each session are required to train a stable model for a BCI-based stroke rehabilitation application.

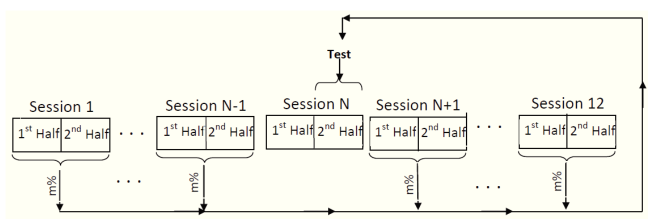

Fig. 2. First experiment: A model is trained using data from 11 sessions and tested on the second half of a new session.

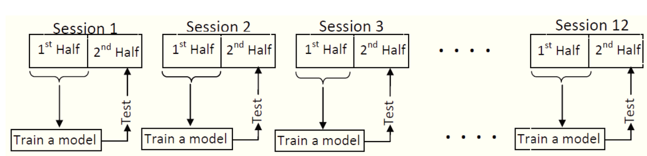

Fig. 3. Second Experiment: A model is trained using the first half of a session and tested on the second half of the same session.

To investigate how well the trained model from the 11 sessions performed on the second half of the new session, the obtained classification accuracies were compared with the classification results of the model trained from the first half of the new session. Thus, in the second experiment, as shown in Fig. 3, the first half of each session was used as the calibration data to train a model.

In all the experiments, the EEG data from 0.5 to 2.5 s after the cue were extracted and filtered using 9 band-pass Chebyshev Type II. Then, the CSP filtering was performed in each band, and a reduced set of features from all the bands was selected using the Mutual Information-based Best Individual Feature algorithm [2]. It is noted that in this study, for each applied CSP, two pairs of filters were used, and for all the mentioned algorithms the LDA classifier was applied in the classification step.

IV. RESULTS AND DISCUSSION

This section investigates the possibility of omitting the calibration phase if we have 11 sessions recorded over a month from a same user. In the first step, the second half of each session of a subject was considered as the test data, and all the data from the other 11 sessions were used as the train data. We aimed to know whether a big amount of data recorded over one month can capture most of the variabilities in the data and make a model working well on the data recorded in a new session. For comparison purpose, in the second step, for each session, the first half of the data were used for training a model and subsequently the second half of the data were used for testing. Since in this step the train (calibration) and test data were recorded one after the other in a same day, we expected to have less variabilities between the train (calibration) and the test data. Thus, the classification accuracies of these models are considered as the baseline, and compared with the classification accuracies obtained by the models trained from the other 11 sessions (see Fig. 4).

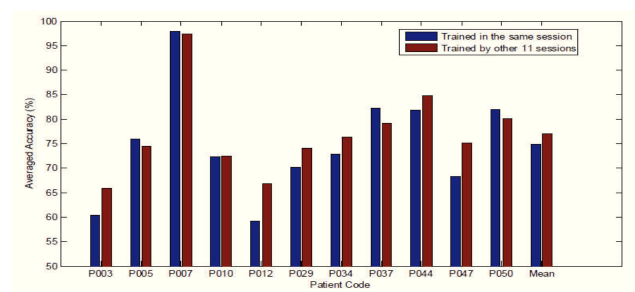

Fig. 4. Comparing performances of the models trained by different train data. Blue bars show the classification accuracies obtained by the models trained using the data recorded directly before the test data. The red bars show the classification accuracies obtained by the models trained using 11 sessions data recorded in different days rather than the test day.

Fig. 4 shows that using $2 * 160 * 11 = 3432$ trials from other sessions recorded over one month can provide a good model which is stable against the variabilities in the new session. Interestingly, the models trained from the 11 sessions outperformed the models trained from the data recorded instantly before the test session by an average of 2.1%. More interestingly, statistical t-test on all the classification results showed that for 4 subjects (P003, P034, P047, P050) the model obtained by the data from 11 sessions significantly performed better than the model obtained from the data recorded just before the test data. For the 7 remaining subjects there is no statistically difference between the results of these two groups of the models.

The results presented at Fig. 4 suggested that there is no need to have a specific calibration phase for a long-term BCI user performing stroke rehabilitation if we have large enough data recorded previously from the same user. In our experiment, each session had 160 trials from each class. So this question arises whether a reliable model still can be trained, if we have less number of trials in each session. In the other words, how many trials over 11 sessions are required to train a stable and reliable model for stroke rehabilitation?

To answer this question, for each subject we randomly and equally selected different number of trials from each of 11 sessions to train a model, and subsequently we tested the model on the second half of the new session. The averaged results over all the sessions and all the 11 patients were displayed in Fig. 5.

Fig. 5 shows that at the beginning, increasing the number of trials dramatically increased the accuracy. However, after applying around 400 trials of each class the slop of increasing the accuracy was considerably reduced. This means that the trained model is getting stable against increasing the number of the train trials. The thick solid line in Fig. 5 presents the averaged accuracy obtained by the models trained from the data recorded in the same session as the test data. Thus, this figure suggested that on average the models trained using around 660 trials (60 trials from each session) per each class can not only outperform the models trained using the data recorded in the same session as the test data, but also perform only less than two percents less accurate than a model trained using all the $160 * 11 = 1760$ trials per class.

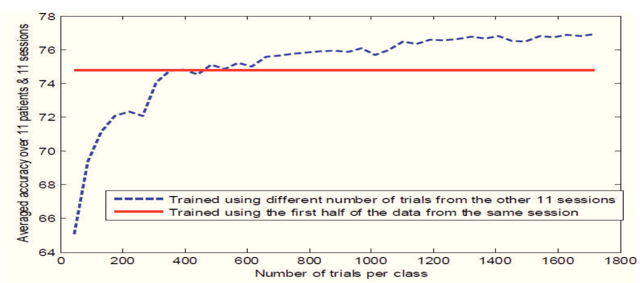

Fig. 5. The averaged accuracy of models trained using different numbers of trials selected from the 11 sessions recorded in the different days rather than the test day. As the baseline, the red line presents the averaged accuracy obtained by the models trained from the data recorded directly before the test data.

V. CONCLUSIONS

Using a large EEG dataset recorded from 11 stroke patients performing 12 stroke rehabilitation sessions over a month, this paper showed that the calibration session can be reliably omitted for long-term BCI users. The experimental results showed that concatenating 11 sessions which each session comprises only 60 trials per class yielded a reliable train model, where to classify a new test session, this model outperformed the standard calibration model recorded directly before the test session by an average of 0.22%. Increasing the number of trials gradually increased the accuracy such that using 11 sessions with 160 trials per class outperformed the standard calibration model by an average of 2.1%. Interestingly, for 4 of 11 subjects, the proposed model significantly outperformed the standard calibration model ($p < 0.05$).

In this study we assumed that for each subject a pool of sessions recorded previously in the same conditions as the new session is available. Since BCI will become more popular in stroke rehabilitation, such pools of sessions will be naturally obtained, and our study would help to make use of this data to facilitate the further sessions. In the future, we aim to further optimize our results by finding the minimum number of sessions required to reliably omit the intra-session calibration.

REFERENCES

[1] N. Birbaumer, "Brain-computer-interface research: Coming of age," *Clin. Neurophysiol.*, vol. 117, pp. 479- 483, 2006.

[2] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, pp. 767-791, 2002.

[3] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang et al, "A Large Clinical Study on the Ability of Stroke Patients in Using EEG-Based Motor Imagery Brain-Computer Interface", *Clinical EEG and Neuroscience*, vol. 42, no. 4, pp. 253-258,Oct. 2011.

[4] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, "EEG-based discrimination between imagination of right and left hand movement," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 6, pp. 642-651, 1997.

[5] M. Krauledat, M. Schroder, B. Blankertz, and K. Muller, Reducing calibration time for brain-computer interfaces: A clustering approach, *Advances Neural Information Processing Syst.*, vol. 19, pp. 753, 2007.

[6] M. Krauledat, M. Tangermann, B. Blankertz, K. Muller, "Towards zero training for Brain-Computer interfacing," *PLoS ONE*, vol. 3, 2008).

[7] S. Fazli, F. Popescu, M. Danoczy, B. Blankertz, K. Muller, C., Grozea, "Subjectindependent mental state classification in single trials", *Neural Networks*, vol. 22, pp. 1305-12, 2009.

[8] C. Calautti, F. Leroy, J. Y. Guincestre, and J. C. Baron, Dynamics of motor network overactivation after striatocapsular stroke: A longitudinal PET study using a fixed-performance paradigm, *Stroke*, vol. 32, pp. 25342542, 2001.

[9] I. Loubinoux, C. Carel, J. Pariente, S. Dechaumont et al, Correlation between cerebral reorganization andmotor recovery after subcortical infarcts, *NeuroImage*, vol. 20, pp. 21662180, 2003.

[10] G. Nelles, G. Spiekermann, M. Jeuptner et al, Reorganization of sensory and motor systems in hemiplegic stroke patients: A positron emission tomography study, *Stroke*, vol. 30, pp. 15101510, 1999.

[11] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang, C. Kuah, C. Wang, K. S. Phua, Z. Y. Chin, and H. Zhang, "Clinical study of neurorehabilitation in stroke using EEG-based motor imagery brain-computer interface with robotic feedback," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2010, pp. 5549-5552.

[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.*, Hong Kong, 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 Recognition*, vol. In Press, Corrected Proof, 2011.

[14] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 41-56, 2008.

[15] M. Pregenzer and G. Pfurtscheller, "Frequency component selection for an EEG-based brain to computer interface," *IEEE Transactions on Rehabilitation Engineering*, vol. 7, pp. 413-419, 1999.

[16] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 4, pp. 441-446, Dec. 2000.