CLASSIFICATION OF KINETICS OF MOVEMENT FOR LOWER LIMB USING COVARIATE SHIFT METHOD FOR BRAIN COMPUTER INTERFACE

A. Hassan[1*], *I. Niazi*[2†], *M. Jochumsen*[2], *F. Riaz*[1], *K. Dremstrup*[2]

[1] College of Electrical and Mechanical Engineering, National University of Sciences and Technology (NUST), Islamabad, Pakistan.
[2] Department of Health Science and Technology (HST), Aalborg University, Denmark.

ABSTRACT

Detecting movement intentions from Electroencephalography (EEG) signals and extracting intended kinetic information such as force and speed may have implications for rehabilitation with assistive technologies by casually linking afferent feedback from the assistive device with the cortical generated movement potentials. However, extraction and classification of kinetics from the 'movement intention' (*before task onset*) on a single-trial basis have only been performed with limited performance due to low signal-to-noise ratio and large trial-to-trial variability. The aim of this study was to investigate a covariate shift method to address the basic challenge of non-stationarity (changes from session to session and trial-to-trial variability) for decoding different levels of speed and force. We tested this method using cross-validation procedures and a linear support vector machine to classify temporal features associated with two levels of force and speed in 9 subjects. The classification accuracy obtained across different class pairs across subjects was 73.1 ± 6.8 % and 70.0 ± 3.6 % with and without the covariate shift method, respectively. The classification accuracy was significantly higher ($p < 0.03$) using the covariate shift method.

***Index Terms*—** Covariate shift, movement-related cortical potentials, brain-computer interface, speed, force.

1. INTRODUCTION

In the last decade brain-controlled devices, commonly known as brain-computer interfaces (BCIs), have emerged as an augmentative tool for rehabilitation to improve or restore the ability of movement-impaired individuals [1]. It has been reported that coincident activation of the brain from somatosensory feedback (from a single pulse electrical stimulation) and motor imagination can promote plastic changes associated with those seen in motor relearning [2] for different patient groups e.g. stroke. This idea was implemented as a BCI by detecting the movement intention of movement-related cortical potential (MRCP) that triggered electrical stimulation [3]. Instead of triggering a single pulse electrical stimulation, the detection of the movement intention can trigger functional electrical stimulation or a robotic device that can replicate a certain movement. However, to replicate the intended movement information about speed and force must be extracted before task onset, and in this way correct afferent feedback can be provided by the assistive device according to the movement intention. This will close the motor control loop and provide the BCI with more degrees of freedom that can be used to introduce variable training that can maximize the retention of re-learned movements during rehabilitation of stroke patients [4]. Extracting different levels of speed and force from movements intentions have been done previously e.g. [5, 6, 7] by using the marginal distribution of the discrete wavelet transform and from temporal features. These were classified with support vector machines (SVMs) with optimized Gaussian kernels.

The extraction of the kinetic information from a limb is impeded by the inherent problem of non-stationarity in EEG signals. This is a big hindrance in the robustness of BCI systems which can be used for rehabilitation. This problem even becomes more challenging if a less number of electrodes is used, which would be desirable for a quick EEG setup in a clinical setting. The factors contributing to the non-stationarity of EEG data are changes in user attention level during sessions, fatigue and differences in the impedance of electrode due to their positioning. Hence, the EEG distributions change from one session to another and even within a single session illustrating the non-stationarity of the data [8].

However, in standard supervised machine learning algorithms the prior probability ($p(x)$) is assumed to remain the same during the training and testing phase. The situation where the prior probability changes between training and testing is called *covariate shift* [9]. In this situation, an adaptive BCI system can be designed in a way, so that it can adjust itself in case there is a shift in the data. In a previous study, the covariate shift has been applied for a two-class problem (discrimination between a movement or an idle state) using

*Thanks to National University of Sciences and Technology, Pakistan for funding this work.
†This work was supported by the Danish Agency for Science and Technology and Innovation, Denmark.

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

event related synchronization in the beta band which appear after termination of the movement [10]. In this paper, only time domain features from the movement intention are extracted to classify force and speed into different levels. The features are classified with an SVM classifier with a *linear* kernel. The aim of this study was to compare the performance of the SVMs with and without the use the covariate shift method.

2. METHODOLOGY

2.1. Subjects

Nine healthy subjects (1 female and 8 males 29 $\pm$ 6 years old) participated in this study. A informed consent was taken from all the subjects before participation. The procedures were approved by the local ethical committee (N-20100067).

2.2. Experimental Protocol

Subjects were seated in a chair with their right foot fixated to a pedal where a force transducer was attached. They were instructed to perform maximum voluntary contractions (MVCs) followed by four different tasks of real isometric dorsi-flexions of the right ankle: i) 0.5 s to reach 20 % MVC (f20) ii) 0.5 s to reach 60 % MVC (f60), iii) 3 s to reach 20 % MVC (s20) and iv) 3 s to reach 60 % MVC (s60). A custom made program called *Follow Me by Knud Larsen, Aalborg University*, was used to assist the subjects cues to perform the movements with correct level of force and speed by providing cues. The subjects were instructed to follow a particular force trace (see Fig: 1). They received visual feedback of their performance during the experiment. Each of the four tasks was repeated 50 times in blocks which were randomized.

2.3. Signal Acquisition

EEG

Only ten channels of monopolar EEG were continuously recorded using scalp electrodes with a sampling rate of 500Hz. The 20 mm Blue Sensor Ag/AgCl, AMBU A/S, Denmark electrodes were placed on the scalp according to the international 10-20 system at FP1, F3, F4, Fz, C3, C4, Cz, P3, P4 and Pz locations. The reference and ground electrodes were placed on the right earlobe and at nasion, respectively. The EEG was divided into epochs using a trigger that was sent from *Follow Me* at the beginning of each trial (at the beginning of the preparation phase in Fig. 1) to the EEG amplifier. FP1 was used to record EOG activity.

Force and MVC

The recordings for force were made using *Mr. Kick* (Knud Larsen, SMI, Aalborg University) and used as input to *Follow Me*. The force was sampled at 2000 Hz. A total of three MVCs were recorded with a rest of one minute from one contraction to another. The MVC having the highest value was used.

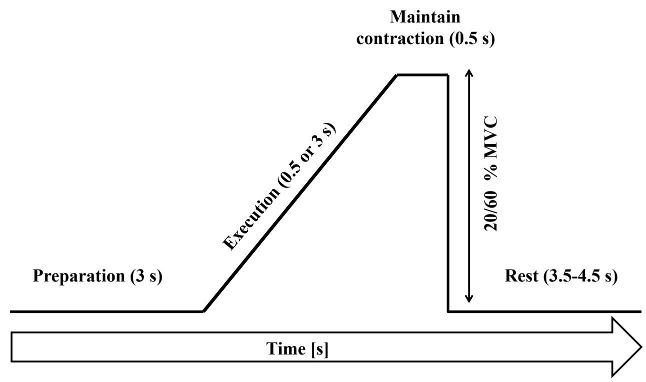

Fig. 1. The subjects had to prepare for 3 s after which they started the execution of the movement. The execution phase lasted either 0.5 or 3 s. When they reached the desired force level (20 or 60 % MVC) they maintained the contraction for 0.5 s followed by a rest period.

Pre-processing

A 2nd order Buttorworth filter was used to bandpass filter the EEG signal from 0.05 to 10 Hz and spatially filtered with a Large Laplacian spatial filter.

2.4. Classification of Movements

Feature extraction

Epochs having EOG activity greater than 125 μV were rejected ($\approx$6 per task). The single-trial movement intentions (2000–100 ms before the task onset) were used to extract six temporal features to predict the intended movement type. The following features were extracted: i) point of max. negativity, ii) mean value of amplitude, iii) and iv) slope and intersection of a linear regression using data from the entire interval, v) and vi) slope and intersection between 500–100 ms of a linear regression before the task onset (see Fig. 2).

Covariate shift

Under the conditions of covariate shift, the prior probability $p(x)$ changes between the training and testing data where the data is assumed to be generated by the model $p(y|x)p(x)$. In this situation, the training data will not accurately represent the statistics of the test data. In the case of covariate shift in

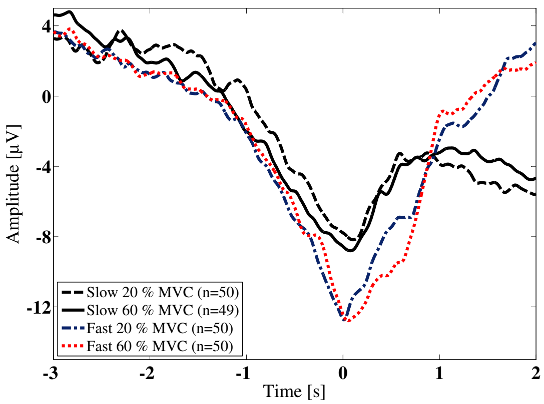

Fig. 2. Average of the four tasks for subject 1. 'Fast' and 'Slow' refers to 0.5 s and 0.3 s respectively, to reach the desired level of MVC, where '0' is the task onset and 'n' is the number of trials in the average.

the data, we would like to assign more weight to those training samples that give the most information about the test data. The goal is to estimate importance weights (IWs) from training (x_i^{tr}) and test samples (x_i^{te}) such that:

$$\beta(x) = p_{te}(x)/p_{tr}(x) \tag{1}$$

where $\beta(x)$ are the IWs. By using these IWs to weight the training data, we push the learning algorithm towards those training samples that give best representation of the test data.

However, estimating the probability distributions for multidimensional data is always a challenge. In this paper, we propose to use Gaussian kernels to transform the input data into high dimensional space. Kernel mean matching (KMM) method proposed by A. Gretton et.al [11] is used to minimise the difference between the training and testing data distributions in the transformed domain. The objective function (J) is given by the difference of the two empirical means of training (n^{tr}) and testing (n^{te}) data samples:

$$J(\beta) = \min_{\beta} \left|\left| \frac{1}{n_{tr}} \sum_{i=1}^{n_{tr}} \beta_i \Phi(x_i^{tr}) - \frac{1}{n_{te}} \sum_{i=1}^{n_{te}} \Phi(x_i^{te}) \right|\right|^2 \tag{2}$$

$$\text{subject to} \quad \beta_i \in [0, B] \quad \text{and} \quad \left| \frac{1}{n_{tr}} \sum_{i=1}^{n_{tr}} \beta_i - 1 \right| \leq \epsilon,$$

where $B > 0$ and $\epsilon > 0$ are tuning parameters and Φ is the Gaussian basis function. The objective function can be expanded into a quadratic programming problem to find suitable β as:

$$\min_{\beta} \quad \frac{1}{2}\boldsymbol{\beta}^T \mathbf{K} \boldsymbol{\beta} - \boldsymbol{\kappa}^T \boldsymbol{\beta}, \tag{3}$$

$$\text{subject to} \quad \beta_i \in [0, B] \text{ and } |\textstyle\sum_{i=1}^{n_{tr}} \beta_i - n_{tr}| \leq n_{tr}\epsilon,$$

where $\mathbf{K}_{ij} = k\left(x_i^{tr}, x_j^{tr}\right)$ and $\boldsymbol{\kappa}_i = \frac{n_{tr}}{n_{te}} \sum_{j=1}^{n_{te}} k\left(x_i^{tr}, x_j^{te}\right)$ and k is the Gaussian kernel mapping function. In this paper, we have used the following values of Gaussian kernel width $\sigma = 0.1$, $B = 1000$, and $\epsilon = (\sqrt{n_{tr}} - 1/\sqrt{n_{tr}})$ for calculating the IWs.

Importance Weighted Support Vector Machines

After calculating the IWs, the next step is to incorporate them into a machine learning algorithm. For our problem, we have used the standard SVMs as the machine learning algorithm because on the complexity and performance scale, they are considered to be optimal. The standard form of soft margin SVMs is:

$$\max_{\alpha} \quad \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\boldsymbol{\alpha}^T \boldsymbol{H} \boldsymbol{\alpha} \tag{4}$$

$$\text{such that} \quad 0 \leq \alpha_i \leq C \text{ and } \textstyle\sum_{i=1}^{n} \alpha_i y_i = 0,$$

where α_i are the Lagrange multipliers, $H_{i,j} \equiv y_i y_j \langle x_i, x_j \rangle$ and C is the penalty factor for misclassification for n number of training samples. By incorporating the IWs into the SVM, the importance weighted SVM (IWSVM) looks like:

$$\max_{\alpha} \quad \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\boldsymbol{\alpha}^T \boldsymbol{H} \boldsymbol{\alpha} \tag{5}$$

$$\text{such that} \quad 0 \leq \alpha_i \leq \beta_i C \text{ and } \textstyle\sum_{i=1}^{n} \alpha_i y_i = 0.$$

For complete mathematical and algorithmic details of IWSVMs, readers are referred to [12]. The only difference between the standard SVM and IWSVM is the limit on the Lagrange multipliers (α_i). This IWSVM is going to adapt the decision boundary to consider the IWs calculated from the training and testing data. This type of SVMs should perform better than a standard SVMs if there exists a data shift in the input data.

3. RESULTS

The 3-fold cross-validation was used to calculate the classification accuracy of each task pair(fast 20% MVC (f20) versus fast 60% MVC (f60) etc.). The results are presented in Fig. 3. The best performance was obtained when the linear SVM was used in combination with covariate shift methodology for f60-s60 class pair ($79.3 \pm 8.7\%$). On average across the subjects and the six task pairs linear SVM with covariate shift methods gave classification accuracy of $73.1 \pm 6.8\%$), whereas linear SVM without the covariate shift method gave $70.0 \pm 3.6\%$ of classification accuracy. A two-way analysis of variance (ANOVA) with factors methods (linear SVM with and without covariate shift) and six class pairs (f20-f60, f20-s20, f20-s60, f60-s20, f60-s60, and s20-s60) showed significance improvement ($F(1, 8) = 6.27, p < 0.03$).

4. DISCUSSIONS

In this paper, we have proposed to use covariate shift to deal with the non-stationarity of EEG signals from one session to

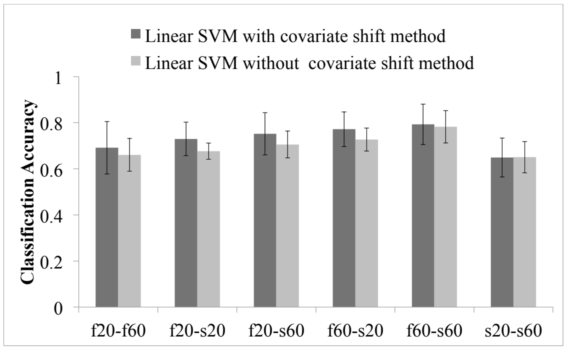

Fig. 3. The columns represents the average (across subjects) classification performance of the SVMs with (dark shade) and without (light shade) covariate shift method for all combinations of the task pairs. The standard deviations are indicated in the figure.

another to classify different movement kinetics before the onset of the task. Any BCI system capable of classifying movement kinetics, with short latency with respect to the onset of a task can be used as a neuromodulatory system [2] and could potentially be used to drive an external augmentative device (robot assisted movements) for restoring or improving the motor functions for stroke patients. As proposed by Mrachacz-Kersting et. al. [2], plasticity can be induced within the human motor cortex if an artificially generated signal (peripheral electrical stimulation) can coincide with a physiologically generated signal (movement intention). The experimental results presented in this paper show that by using a very simple classifier (linear kernel SVM), we can deal with the non-stationarity of EEG signals achieving significant improvement ($p < 0.03$) in classification accuracy over a standard method. The classification accuracies of self-paced neuromodulatory BCI systems have been found to be positively correlated with excitability of the corticospinal projections of the target muscle [3].

It is a known fact that the neural signals generated in self-paced scenarios are different compared to those obtained in cue based scenarios [13]. In this paper, the level of force and speed was precisely controlled across subjects using a cue based system. One of the issues with KMM methodology used in this paper is that it requires the testing data to be available before the IWs can be calculated. Therefore, data recorded from previous sessions can be used to train the IWSVMs. In future, we would like to make covariate shift methods incremental and adaptive to make them directly applicable to online BCI systems.

5. REFERENCES

[1] J. Daly and J. Wolpaw, "Brain-computer interfaces in neurological rehabilitation," *The Lancet Neurology*, vol. 7, no. 11, pp. 1032–1043, 2008.

[2] N. Mrachacz-Kersting, S. Kristensen, I. Niazi, and D. Farina, "Precise temporal association between cortical potentials evoked by motor imagination and afference induces cortical plasticity," *Journal of Physiology*, vol. 590, no. 7, pp. 1669–1682, 2012.

[3] I. Niazi, N. Mrachacz-Kersting, N. Jiang, K. Dremstrup, and D. Farina, "Peripheral electrical stimulation triggered by self-paced detection of motor intention enhances motor evoked potentials," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 4, pp. 595–604, 2012.

[4] J. Krakauer, "Motor learning: its relevance to stroke recovery and neurorehabilitation," *Current Opinion in Neurology*, vol. 19, no. 1, pp. 84–90, 2006.

[5] D. Farina, O. Nascimento, M. Lucas, and C. Doncarli, "Optimization of wavelets for classification of movement-related cortical potentials generated by variation of force-related parameters," *Journal of Neuroscience Methods*, vol. 162, no. 1-2, pp. 357–363, 2007.

[6] Y. Gu, O. Nascimento, M. Lucas, and D. Farina, "Identification of task parameters from movement-related cortical potentials," *Medical & Biological Engineering & Computing*, vol. 47, no. 12, pp. 1257–1264, 2009.

[7] M. Jochumsen, I. Niazi, N. Mrachacz-Kersting, D. Farina, and K. Dremstrup, "Detection and classification of movement-related cortical potentials associated with task force and speed," *Journal of Neural Engineering*, vol. 10, no. 5, pp. 056015, 2013.

[8] Y. Li, H. Kambara, Y. Koike, and M. Sugiyama, "Application of covariate shift adaptation techniques in brain computer interfaces," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp. 1318–1324, 2010.

[9] J. Quiñonero-Candela, M. Sugiyama, A. Schwaighofer, and N. Lawrence, Eds., *Dataset Shift in Machine Learning*, MIT Press, Cambridge, MA, 2009.

[10] R. Mohammadi, A. Mahloojifar, and D. Coyle, "Unsupervised short-term covariate shift minimization for self-paced bci," in *IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain*, Singapore, 2013.

[11] A. Gretton, A. Smola, J. Huang, M. Schmittfull, K. Borgwardt, and B. Schoelkopf, *Dataset Shift in Machine Learning*, chapter 8. Covariate Shift and Local

Learning by Distribution Matching, pp. 131–160, MIT Press, Cambridge, MA, 2009.

[12] A. Hassan, R. Damper, and M. Niranjan, "On acoustic emotion recognition: Compensating for covariate shift," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 21, no. 7, pp. 1458–1468, 2013.

[13] S. Jankelowitz and J. Colebatch, "Movement–related potentials associated with self-paced, cued and imagined arm movements," *Experimental Brain Research*, vol. 147, no. 1, pp. 98–107, 2002.