Towards a Cure for BCI Illiteracy

Carmen Vidaurre[a], Benjamin Blankertz[a,b]
[a]Berlin Institute of Technology, Machine Learning Laboratory, Berlin, Germany
[b]Fraunhofer FIRST IDA, Berlin, Germany

Correspondence: Carmen Vidaurre, TU-Berlin, Sekr. FR 6-9, Franklinstr. 28/29, 10587 Berlin, Germany. E-mail: vidcarATcsDOTtu-berlinDOTde, phone +49 3031478625, fax+49 3031478622

Abstract. Brain-Computer Interfaces (BCIs) allow a user to control a computer application by brain activity as acquired, e.g., by EEG. One of the biggest challenges in BCI research is to understand and solve the problem of 'BCI Illiteracy', which is that BCI control does not work for a non-negligible portion of subjects (estimated 15% to 30%). Here, we investigate the illiteracy problem in BCI systems which are based on the modulation of sensorimotor rhythms. In this paper, a sophisticated adaptation scheme is presented which guides the user from an initial subject-independent classifier that operates on simple features to a subject-optimized state-of-the-art classifier within one session while the user interacts the whole time with the same feedback application and does not need to care about what is going on behind the scenes. While initial runs use supervised adaptation methods for robust co-adaptive learning of user and machine, final runs use unsupervised adaptation and therefore provide an unbiased measure of BCI performance. Using this approach, which does not involve any offline calibration measurement, good performance was obtained by good BCI subjects (also one novice) after 3-6 minutes of adaptation. More importantly, the use of machine learning techniques allowed subjects who were unable to achieve successful feedback before to gain significant control over the BCI system. In particular, one subject had no peak of the sensory motor idle rhythm in the beginning of the experiment, but could develop such peak during the course of the session (and use voluntary modulation of its amplitude to control the feedback application).

Keywords: Brain Computer Interface (BCI); Sensorimotor Rhythm (SMR); BCI Illiteracy; Co-adaptive Learning, Machine Learning.

1. Introduction

Brain-Computer Interfaces (BCIs) systems aim to provide users control over a computer application by their brain activity. In EEG-based BCIs, one of the biggest research challenges is to understand and solve the problem of 'BCI Illiteracy', which is that BCI control does not work for a non-negligible portion of subjects (estimated 15% to 30%). In a screening study, N=80 subjects performed motor imagery first in a calibration (i.e. without feedback) measurement and then in a feedback measurement in which they could control a 1D cursor application. Coarsely, we observed three categories of subjects: subjects for whom (I) a classifier could be successfully trained and who performed feedback with good accuracy; (II) a classifier could be successfully trained, but feedback did not work well; (III) no classifier with acceptable accuracy could be trained. Whereas subjects of Cat. II had obviously difficulties with the transition from offline on online operation, subjects of Cat. III did not show the expected modulation of sensorimotor rhythms (SMRs): either no SMR idle rhythm was observed over motor areas, or this idle rhythm was not attenuated during motor imagery.

Here we present preliminary results of a one-session pilot study in which it was investigated, whether co-adaptive learning using machine-learning techniques could help subjects of Cat. II and III to achieve successful feedback. Our results show that adaptive machine learning methods successfully helped subjects who suffered from the BCI illiteracy problem before, to gain control of the system.

2. Material and Methods

2.1. Experimental setup

The study consisted of a one-day session that immediately started with BCI feedback using a pre-trained subject-independent classifier, as in [Vidaurre et al., 2007]. Using supervised and unsupervised techniques, the classifier was adapted to the specific brain signals of the experimental subject during the session. Adaptation was performed during three incremental levels. While the feedback application

itself stayed the same for the whole experiment, the features on which the classifier operated and the adaptation methods changed from level to level as described below.

2.2. Methods

Eleven subjects took part in the study. Six subjects belonged to Cat. I (for one novice subject, no prior data was available, but she turned out to be a Cat. I subject), two further subjects belonged to Cat. II and three subjects to Cat. III. All subjects performed 7 feedback runs, each of them consisting of 100 trials (50 trials of each class). Feedback was provided in form of a moving cursor, whose speed was determined by the classification output (similar to [Blankertz et al., 2008b]). The task of the subject was to use motor imagery to make the cursor move into a previously indicated target direction. Two different types of motor imagery, chosen out of three possibilities (motor imagery of left hand, right hand or foot) were selected in advance. For seven subjects, previous data with motor imagery performance was available which revealed which two motor imagery tasks should be used. For the other four subjects (three of Cat. III and one novice) no prior information could be used and they were asked to select two out of the three possible motor imagery tasks.

Throughout the whole session, all classifiers were based on Linear Discriminant Analysis (LDA). When advisable due to high dimensionality of features, the estimation of the covariance matrix that is needed for LDA was corrected by shrinkage ([Ledoit and Wolf, 2004]). In order to define the adaptation schemes for LDA we use a specific variant that is introduced here. For LDA the covariance matrices of both classes are assumed to be equal (assumption of linear separability) and it will be denote by Σ here. Furthermore we denote the means of the two classes by μ_1 and μ_2, an arbitrary feature vector by $\mathbf{x}$ and define:

$$D(\mathbf{x}) = [b; \mathbf{w}]^{\mathrm{T}} \cdot [1; \mathbf{x}] \tag{1}$$

$$\mathbf{w} = \Sigma^{-1} \cdot (\mu_2 - \mu_1) \tag{2}$$

$$b = -\mathbf{w}^{\mathrm{T}} \cdot \mu \tag{3}$$

$$\mu = (\mu_1 + \mu_2) / 2 \tag{4}$$

where $D(\mathbf{x})$ is the difference in the distance of the feature vector $\mathbf{x}$ to the separating hyperplane, which is described by its normal vector $\mathbf{w}$ and bias b. Usually, the covariance matrix used in Eq. (2) is the class-average covariance matrix. But it can be shown that using the pooled covariance matrix (which can be estimated without using label information, just by aggregating the features of all classes) yields the same separating hyperplane. In this study we used the pooled covariance matrix in Eq. (2). Similarly, the class-average mean (calculated in Eq. (4)) can be replaced by the pooled mean (average over all feature vectors of all classes). This implies that the bias of the separating hyperplane can be estimated (and adapted) in an unsupervised manner (without label information). The restriction of the method is to have an estimate of the prior probabilities of the 2 classes.

If LDA is to be used as a classifier, observation $\mathbf{x}$ is classified as class 1, if $D(\mathbf{x})$ is less than 0, and otherwise as class 2. But in the cursor control application we use the classifier output $D(\mathbf{x})$ as real number to determine the speed of the cursor.

Finally, we introduce the features and classifiers that have been used in the three levels of the experiment, including three on-line adaptation schemes: the first two are supervised, i.e., they require information about the class label (type of motor imagery task) of the past trial in order to update the classifier. The last method updates the classifier without knowing the task of the past trial (unsupervised adaptation).

Methods for level 1(runs 1-3)

The first run started with a pre-trained subject-independent classifier on simple features: band-power in alpha (8-15 Hz) and beta (16-32 Hz) frequency range in 3 Laplacian channels at C3, Cz, C4. During these runs, the LDA classifier was adapted to the subject after each trial. The inverse of the pooled covariance matrix (see Eq. (2)) was updated for observation $\mathbf{x}(t)$ using a recursive-least-square algorithm [Vidaurre et al., 2006]:

$$\Sigma(t) = \frac{1}{(1 - UC)} \left(\Sigma(t-1) - \frac{\mathbf{v}(t) \cdot \mathbf{v}^T(t)}{\frac{1 - UC}{UC} + \mathbf{x}^T(t) \cdot \mathbf{v}(t)} \right) \tag{5}$$

where $\mathbf{v}(t) = \Sigma^{-1}(t-1)\cdot\mathbf{x}(t)$. Note, the term $\mathbf{x}^T(t)\cdot\mathbf{v}(t)$ is a scalar and no costly matrix inversion is needed. In Eq. (5), *UC* stands for update coefficient and is a small number between 0 and 1. For the present study, we chose *UC*= 0.015 based on a simulation using the data of the screening study. To estimate the class-specific adaptive mean $\boldsymbol{\mu}_1(t)$ and $\boldsymbol{\mu}_2(t)$ one can use an exponential moving average:

$$\boldsymbol{\mu}_i(t) = (1-UC)\cdot\boldsymbol{\mu}_i(t-1) + UC\cdot\mathbf{x}(t) \tag{6}$$

where *i* is the class of $\mathbf{x}(t)$ and *UC* was chosen to be 0.05. Note that the class-mean estimation is done in a supervised manner.

Methods for level 2 (runs 4-6)

For the subsequent 3 runs, a classifier was trained on a more complex composed band-power feature. On the data of run 1-3, a subject-specific narrow band was chosen automatically ([Blankertz et al., 2008b]). For this frequency band, optimized spatial filters have been determined by Common Spatial Pattern (CSP) analysis ([Blankertz et al., 2008b]). Furthermore, six Laplacian channels have been selected according to their discriminability, which was quantified by a robust variant of the Fisher score (mean replaced by median). The selection of the positions was constraint such that two positions have been selected from each of the areas over left hand, right hand and foot. While CSP filters were static, the position of the Laplacians was reselected based on the Fisher score of the channels. Channel selection and classifier were recalculated after each trial using the last 100 trials. The feature vector was the concatenation of log band-power in the CSP channels and the selected Laplacians channels. The addition of the repeatedly selected Laplacian channels was included in order to provide flexibility with respect to spatial location of modulated brain activity. During these three runs the adaptation to the subject was done again in a supervised way.

Methods for level 3 (runs 7-8)

Finally for the last 2 runs, CSP filters have been calculated on the data of runs 4-6 and a classifier was trained on the resulting log band-power features. The bias of the classifier in Eq. (3) was adapted by updating the pooled mean $\boldsymbol{\mu}$ after each trial with *UC*=0.05. The update rule for the pooled mean was analogue to Eq. (6), but without distinction by class labels. Note that this adaptation scheme is unsupervised. For more information about unsupervised methods, see [Vidaurre et al., 2008].

3. Results

As a verification of the novel experimental design, we first discuss the results for the six subjects of Cat. I. Here, very good feedback performance was obtained within the first run after 20 to 40 trials (i.e., after 3-6 minutes) of adaptation and hit rates increased further in runs 2 and 3 and stayed on that level in subsequent runs, see Fig. 1(a). The challenge of the experiment was the performance with the two subjects of Cat. II and three subjects of Cat. III. All those five subjects did not have control in the first three runs, but they were able to gain it when the machine learning based techniques came into play in runs 4-6: in the average performance for Cat. II a sudden jump occurred from run 3 to run 4, and for Cat. III a continuous increase in runs 4 to 5, see Fig. 1(b). Note that the obtained level of performance could be maintained (and partly even improved) in runs 7 and 8, where adaptation was performed in an unsupervised fashion.

4. Conclusion

Results in Fig. 1(a) show that subjects suffering from BCI illiteracy (i.e., belonging to Cat. II or III) before could gain BCI control within one session using a novel incremental adaptation paradigm. In particular, one subject who had no peak of the SMR idle rhythm in the beginning of the measurement could develop such with our adaptive feedback training, see Fig. 1(b). This important finding gives rise to the development of neurofeedback training procedures that might help to cure BCI illiteracy. Further studies with a larger number of subjects are required in order to confirm these initial findings.
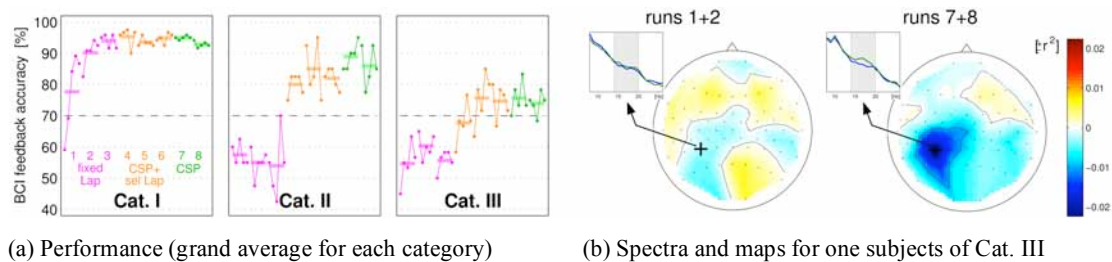
(a) Performance (grand average for each category)   (b) Spectra and maps for one subjects of Cat. III

Figure 1. *(a) Grand average of feedback performance within each run (horizontal bars and dots for each group of 20 trials) for subjects of Cat. I (N=6), Cat. II (N=2) and Cat. III (N=3). An accuracy of 70% is assumed to be a threshold required for BCI applications. Note that all runs of one subject have been recorded within one session. (b) For one subject of Cat. III spectra in channel CP3 and scalp topographies of band-power differences (signed r^2-values) between the motor imagery conditions are compared between the beginning (runs 1+2) and the end (runs 7+8) of the experiment.*

Acknowledgements

References

Blankertz B, Losch F, Krauledat M, Dornhege G, Curio G, Müller KR. The Berlin Brain-Computer Interface: Accurate Performance From First-Session in BCI-Naive Subjects. *IEEE Transactions on Biomedical Engineering*, 55:2452-2462, 2008.

Blankertz B, Tomioka R, Lemm S, Kawanabe M, Müller KR. Optimizing spatial filters for robust EEG single- trial analysis. *IEEE Signal Processing Magazine*, 25:41-56, 2008b.

Ledoit O, Wolf M. A well-conditioned estimator for large-dimensional covariance matrices, Journal of Multivariate Analysis 88(2): 365-411, 2004.

Vidaurre C, Schlögl A, Scherer R, Cabeza R, Pfurtscheller G. A fully on-line adaptive BCI. *IEEE Transactions on Biomedical Engineering,* 53:1214-1219, 2006.

Vidaurre C, Schlögl A, Scherer R, Cabeza R, Pfurtscheller G. Study of on-line adaptive discriminant analysis for EEG-based brain computer interfaces. *IEEE Transactions on Biomedical Engineering,* 54:50-56, 2007.

Vidaurre C, Schlögl A, Blankertz B, Kawanabe M, Müller KR. Unsupervised adaptation of the LDA classifier for Brain-Computer Interfaces. In Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course 2008, Verlag der Technischen Universität Graz, 2008, 122-127.