This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2628057, IEEE Transactions on Neural Systems and Rehabilitation Engineering

Open Access Dataset for EEG+NIRS Single-Trial Classification

Jaeyoung Shin, Alexander von Lühmann, Benjamin Blankertz, Do-Won Kim, Jichai Jeong, Han-Jeong Hwang*
and Klaus-Robert Müller*, *Member, IEEE*

***Abstract*—We provide an open access dataset for hybrid brain-computer interfaces (BCIs) using electroencephalography (EEG) and near-infrared spectroscopy (NIRS). For this, we conducted two BCI experiments (left vs. right hand motor imagery; mental arithmetic vs. resting state). The dataset was validated using baseline signal analysis methods, with which classification performance was evaluated for each modality and a combination of both modalities. As already shown in previous literature, the capability of discriminating different mental states can be enhanced by using a hybrid approach, when comparing to single modality analyses. This makes the provided data highly suitable for hybrid BCI investigations. Since our open access dataset also comprises motion artifacts and physiological data, we expect that it can be used in a wide range of future validation approaches in multimodal BCI research.**

***Index Terms*— Brain-computer interface (BCI), electroencephalography (EEG), hybrid BCI, mental arithmetic, motor imagery, near-infrared spectroscopy (NIRS), open access dataset.**

I. INTRODUCTION

BRAIN-computer interface (BCI) technology has increasingly been receiving attention as an alternative communication option for severely paralyzed patients in the last decades [1-3]. Over the last years, BCIs have also been used as an augmentative tool for rehabilitation [4-6]. Commercial and non-medical BCI applications were developed [7-13], and novel analysis algorithms and signal processing techniques were introduced for developing practical BCIs [14-18]. However, whenever investigating new BCI methods, many experiments have to be conducted for each independent investigation. This consumes much time and results in a lack of reproducibility because the measured dataset exploited for each study is usually not shared or available to the public for reusable purposes. Thus, we identified a growing need to publish an open access dataset for BCI research. So far, several electroencephalography (EEG)-based BCI datasets have been published [19-23]. These datasets were used to evaluate the state-of-the-art of signal processing and classification methods [24-30].

Recent studies verified the superior performance of hybrid BCIs based on EEG and near-infrared spectroscopy (NIRS) compared to that of single modality solutions [31-36], and new customized instrumentation approaches for hybrid BCI have recently been presented [37]. Despite this ongoing trend towards hybrid BCI technology, there is no open access dataset for hybrid BCIs available so far. Thus, our paper aims to provide an open access brain signal dataset suitable for investigating hybrid BCI research. For the sake of generality, we chose two sorts of mental tasks to build the dataset: motor imagery (MI) whose practicability and usefulness were repeatedly demonstrated by numerous numbers of both EEG and NIRS-based BCI studies [38, 39] and mental arithmetic (MA) which is one of the most widely used paradigms in NIRS-based BCI research [40, 41].

In addition, representative motion artifacts induced by ocular and head movements were separately recorded and included in the dataset, but not intensively analyzed in this study. The motion artifact data can be potentially used to develop new algorithms that can improve the signal quality and robustness [42, 43], or explore methods for artifact rejection depending on the research purpose.

There are many challenges and opportunities in the signal analysis of multimodal data for hybrid BCIs. The purpose of this study was the demonstration of the suitability of the open access data. We performed basic decoding procedures and provide the results along with the data. Accuracies can be improved by applying more sophisticated and state-of-the-art methods.

In Section II, we describe details of each dataset, including experimental paradigms, recording methods and analysis methods. The baseline analysis results are presented in Section III, and related discussions are provided in Section IV. The conclusion of this study is given in Section V.

II. MATERIALS AND METHODS

A. Subjects

Twenty-eight right-handed and one left-handed healthy subjects participated in this study (fourteen males and fifteen females, average age (years) 28.5 ± 3.7 (mean ± standard deviation)). None of them reported neurological, psychiatric or other brain-related diseases. All volunteers were informed

This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2014R1A6A3A03057524) and Ministry of Science, ICT & Future Planning (NRF-2015R1C1A1A02037032), and supported in part by the Brain Korea 21 PLUS Program through the NRF funded by the Ministry of Education, the Korea University Grant and BMBF (#01GQ0850, Bernstein Focus: Neurotechnology).

*Corresponding Authors

J. Shin and A. von Lühmann are with Machine Learning Group, the Department of Computer Science, Berlin Institute of Technology, Berlin, Germany.

B. Blankertz is with Neurotechnology Group, Berlin Institute of Technology, Berlin, Germany.

D. Kim is with Machine Learning Group, the Department of Computer Science, Berlin Institute of Technology, Berlin, Germany and also with Institute of Biomedical Engineering, Hanyang University, Seoul, Korea.

J. Jeong is with Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea.

H. Hwang is with the Department of Medical IT Convergence Engineering, Kumoh National Institute of Technology, Kumi, Korea (email: h2j@kumoh.ac.kr).

K.-R. Müller is with Machine Learning Group, the Department of Computer Science, Berlin Institute of Technology, Berlin, Germany, and also with the Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea (email: klaus-robert.mueller@tu-berlin.de).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2628057, IEEE Transactions on Neural Systems and Rehabilitation Engineering

about the experimental procedure and written consent was obtained from all participants. They were financially reimbursed after the experiment. This study was conducted according to the declaration of Helsinki and was approved by the Ethics Committee of the Institute of Psychology and Ergonomics, Technical University of Berlin (approval number: SH_01_20150330).

B. Data Acquisition

EEG and NIRS data was collected in an ordinary bright room. EEG data was recorded by a multichannel BrainAmp EEG amplifier with thirty active electrodes (Brain Products GmbH, Gilching, Germany) with linked mastoids reference at 1000 Hz sampling rate. The EEG amplifier was also used to measure the electrooculogram (EOG), electrocardiogram (ECG) and respiration with a piezo based breathing belt. Thirty EEG electrodes were placed on a custom-made stretchy fabric cap (EASYCAP GmbH, Herrsching am Ammersee, Germany) and placed according to the international 10-5 system (AFp1, AFp2, AFF1h, AFF2h, AFF5h, AFF6h, F3, F4, F7, F8, FCC3h, FCC4h, FCC5h, FCC6h, T7, T8, Cz, CCP3h, CCP4h, CCP5h, CCP6h, Pz, P3, P4, P7, P8, PPO1h, PPO2h, POO1, POO2 and Fz for ground electrode) [19]. NIRS data was collected by NIRScout (NIRx GmbH, Berlin, Germany) at 12.5 Hz sampling rate. Each adjacent source-detector pair creates one physiological NIRS channel. Fourteen sources and sixteen detectors resulting in thirty-six physiological channels were placed at frontal (nine channels around Fp1, Fp2, and Fpz), motor (twelve channels around C3 and C4, respectively) and visual areas (three channels around Oz). The inter-optode distance was 30 mm. NIRS optodes were fixed on the same cap as the EEG electrodes. Ambient light was blocked by a firm contact between NIRS optodes and scalp and use of an additional opaque cap over the stretchy fabric cap. Fig. 1 shows the placement of the EEG electrodes and NIRS optodes. EOG was recorded using two vertical (above and below left eye) and two horizontal (outer canthus of each eye) electrodes. ECG was recorded based on Einthoven triangle derivations I and II, and respiration was measured using a respiration belt on the lower chest. EOG, ECG and respiration were sampled at the same sampling rate of the EEG. ECG and respiration data were not analyzed in this study, but are provided along with the other signals. All signals were recorded simultaneously. In order to synchronize the signals, triggers were sent to each instrument at the same time via parallel port using MATLAB.

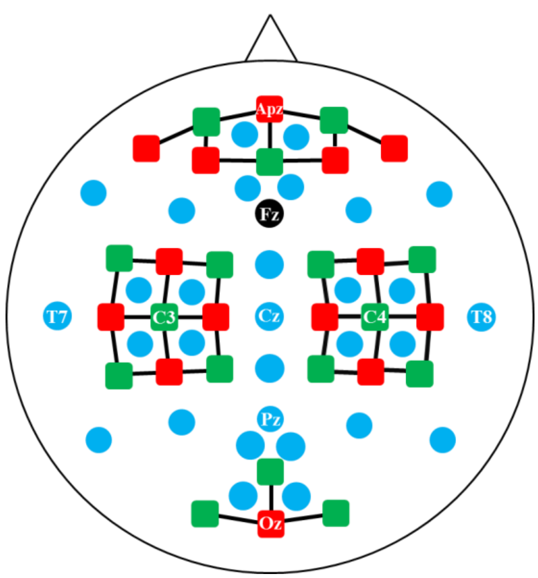

Fig. 1. Placement of EEG electrodes (blue and black (ground) circles) and NIRS sources (red squares) and detectors (green squares). Black solid lines denote NIRS channels.

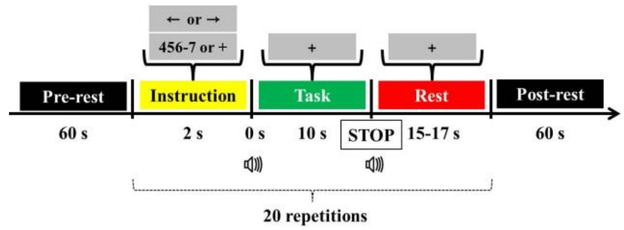

Fig. 2. Schematic sequence diagram of the experimental paradigm. Each session comprised a 1 min pre-experiment resting period, 20 repetitions of the given task and a 1 min post-experiment resting period. The task started with 2 s of a visual introduction of the task, followed by 10 s of a task period and resting period which was given randomly from 15 to 17 s. At the beginning and end of the task period, a short beep (250 ms) was played.

C. Experimental Paradigm

The subjects sat on a comfortable armchair in front of a 50-inch white screen. The distance between their heads and the screen was 1.6 m. They were asked not to move any part of the body during the data recording. The experiment consisted of three sessions of left and right hand MI (dataset A) and MA and baseline tasks (taking a rest without any thought) (dataset B) each. Each session comprised a 1 min pre-experiment resting period, 20 repetitions of the given task and a 1 min post-experiment resting period. The task started with 2 s of a visual introduction of the task, followed by 10 s of a task period and resting period which was given randomly from 15 to 17 s. At the beginning and end of the task period, a short beep (250 ms) was played. All instructions were displayed on the white screen by a video projector. MI and MA tasks were performed in separate sessions but in alternating order (i.e., sessions 1, 3 and 5 for MI (dataset A) and sessions 2, 4 and 6 for MA (dataset B)). Fig. 2 shows the schematic diagram of the experimental paradigm. Five sorts of motion artifacts induced by eye and head movements (dataset C) were measured. The motion artifacts were recorded after all MI and MA task recordings. The experiment did not include the pre- and post-experiment resting

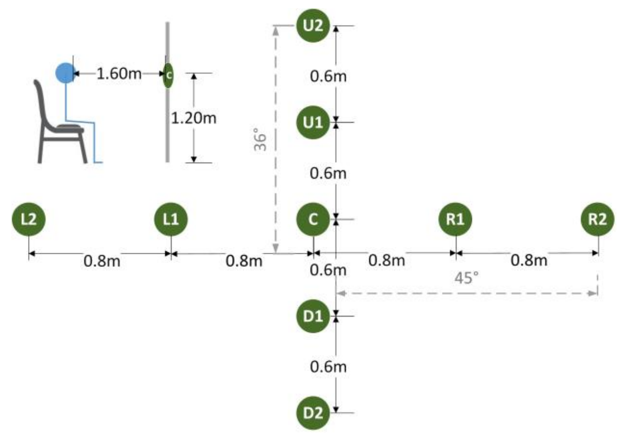

Fig. 3. Set-up for the motion artifact measurements. The characters, ‘C’, ‘L’, ‘R’, ‘U’, and ‘D’, denote center, left, right, up, and down, respectively, and the numbers, 1 and 2, are assigned according to the distance from the center.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2628057, IEEE Transactions on Neural Systems and Rehabilitation Engineering

state periods.

1) Dataset A: MI

For MI, subjects were instructed to perform kinesthetic motor imagery (i.e. to imagine the opening and closing their hands as they were grabbing a ball) to ensure that actual motor imagery, not visual imagery, was performed. All subjects were naive to the MI experiment. For the visual instruction, a black arrow pointing to either the left or right side appeared at the center of the screen for 2 s. The arrow disappeared with a short beep sound and then a black fixation cross was displayed during the task period. The subjects were asked to imagine hand gripping (opening and closing their hands) with a 1 Hz pace. Because there was no exact way to check it during the recording, we repeatedly showed a real hand gripping with a speed of 1 Hz before the recording, thereby leading to performing motor imagery with a constant speed of about 1 Hz as much as possible. Motor imagery was performed continuously over the task period. The task period was finished with a short beep sound and a 'STOP' displayed for 1s on the screen. The fixation cross was displayed again during the rest period and the subjects were asked to gaze at it to minimize their eye movements. This process was repeated twenty times in a single session (10 trials for each left and right hand motor imagery in a single session; 30 trials for each one in the whole three sessions). In a single session, motor imagery tasks were performed on the basis of ten subsequent blocks randomly consisting of one of two conditions: Either first left and then right hand motor imagery or vice versa.

2) Dataset B: MA versus Baseline Task

For the visual instruction of the MA task, an initial subtraction such as 'three-digit number minus one-digit number' (e.g., 384-8) appeared at the center of the screen for 2 s. The subjects were instructed to memorize the numbers while the initial subtraction was displayed on the screen. The initial subtraction disappeared with a short beep sound and a black fixation cross was displayed during the task period in which the subjects were asked to repeatedly perform to subtract the one-digit number from the result of the previous subtraction. For the baseline task, no specific sign but the black fixation cross was displayed on the screen, and the subjects were instructed to take a rest. Note that there were other rest periods between the MA and baseline task periods, as same with the MI paradigm. Both task periods were finished with a short beep sound and a 'STOP' displayed for 1 s on the screen. The fixation cross was displayed again during the rest period. MA and baseline trials were randomized in the same way as MI.

3) Dataset C: Motion Artifacts

We measured five types of motion artifacts. Please refer to a similar study for more details on the experimental set-up [44]. Fig. 3 shows the set-up for the motion artifact measurements in which the characters, 'C', 'L', 'R', 'U', and 'D', denote center, left, right, up and down, respectively, and the numbers, 1 and 2, are assigned according to the distance from the center. Distances between C and L1/R1 and C and L2/R2 markers were 0.8 and 1.6 m, respectively (a visual angle of 45° for L2 and R2from the center) and distances between C and U1 /D1 and C and U2/D2 markers were 0.6 and 1.2 m, respectively (a visual angle of 36° for U2 andD2 from the center).

a) Blinking eyes (BLK)

The subjects were asked to blink their eyes once whenever they heard a short beep sound. It was repeated twenty times in 1 s intervals.

b) Moving eyes (MVE)

The subjects were instructed to move their eyes at 2 s intervals following the arrow indicating the directions to markers on the wall while not moving their head:

- Moving up and down (UD): C→U1→U2→U1→C
- Moving down and up (DU): C→D1→D2→D1→C
- Moving left and right (LR): C→L1→L2→L1→C
- Moving right and left (RL): C→R1→R2→R1→C

This was repeated five times for each direction after a 5 s break.

c) Moving head (MVH)

The subjects moved their heads but without moving their eyes. This was repeated in the same way as *MVE.*

d) Clenching teeth (CLT)

The subjects clenched their teeth strongly for 2 s whenever they heard a short beep sound. This was repeated ten times at 5 s intervals.

e) Opening mouth (OPM)

The subjects opened their mouth widely for 2 s whenever they heard a short beep sound. This was repeated ten times at 5 s intervals.

D. Posterior questionnaire

After completing the experiment, each participant answered a simple posterior questionnaire regarding their impressions and feelings. They rated each question on a 5-point scale (1 point: very low, 5 point: very high) separately for MI and MA. The questionnaire included the following four items:

1) *Concentration: level of concentration during the task*
2) *Fatigue: level of fatigue after the task*
3) *Sleepiness: level of sleepiness or boredom during the task*
4) *Difficulty: level of difficulty of the task*

E. Data processing

1) MI and MA

All data processing was done using MATLAB R2013b (MathWorks, Natick, MA, USA).

The measured EEG data was first re-referenced using a common average reference [45] and filtered (fourth order of Chebyshev type II filter) with a passband of 0.5 - 50 Hz before EOG rejection. Independent component analysis (ICA)-based EOG rejection was performed using the automatic artifact rejection toolbox in EEGLAB [46].

The preprocessed EEG signals were downsampled to 200 Hz. All channels were used for further EEG data processing. Spatial filters were determined by common spatial patterns (CSPs) analysis [14, 47, 48] using the BBCI toolbox [49]. For CSPs of MA data, subject-specific band-pass filter coefficients were estimated by means of the heuristic procedure which estimates a frequency band showing the highest absolute signed squared point biserial correlation coefficient (signed r^2) value [14]. For MI, a subject-independent pass-band from 8 to 25 Hz including μ- and low β-band was selected [50]. For MA, subject-specific pass-bands were selected in the range of 4 - 35 Hz showing the best separability based on the signed r^2. For more details on these bands, please refer to the supporting material. All data processing (i.e. feature extraction and classification) was performed separately on each time window using a moving time window method (window size: 3 s, step size: 1 s). Features were extracted using the log-variance of the first and last three CSP

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2628057, IEEE Transactions on Neural Systems and Rehabilitation Engineering

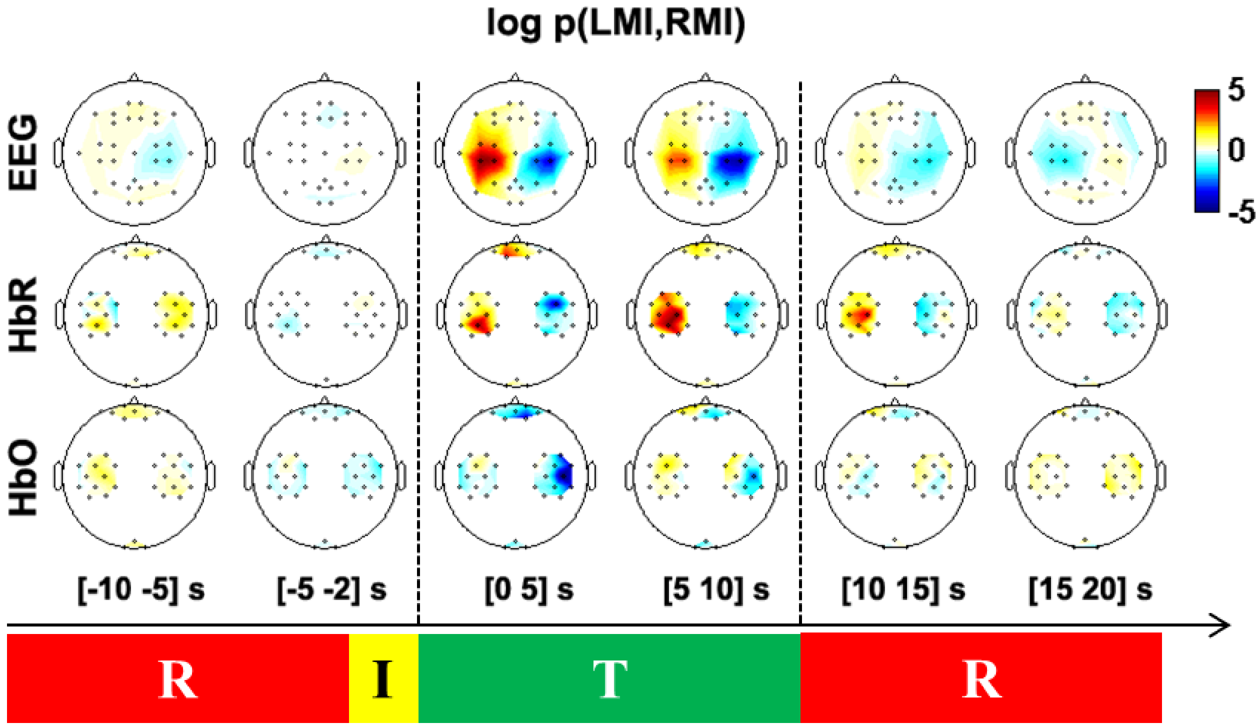

Fig. 4. Scalp evolution of grand average of log (p) based on signed r^2 for motor imagery in EEG and NIRS using the data of 15 subjects showing the best classification performance for better representation. Blue and Red colors denote higher separability of left motor and right imagery, respectively. The numbers displayed denote the time from the task onset. R, I and T denotes rest, instruction and task period, respectively. For better understanding, baseline correction was performed by subtracting the average value between -5 and -2 s.

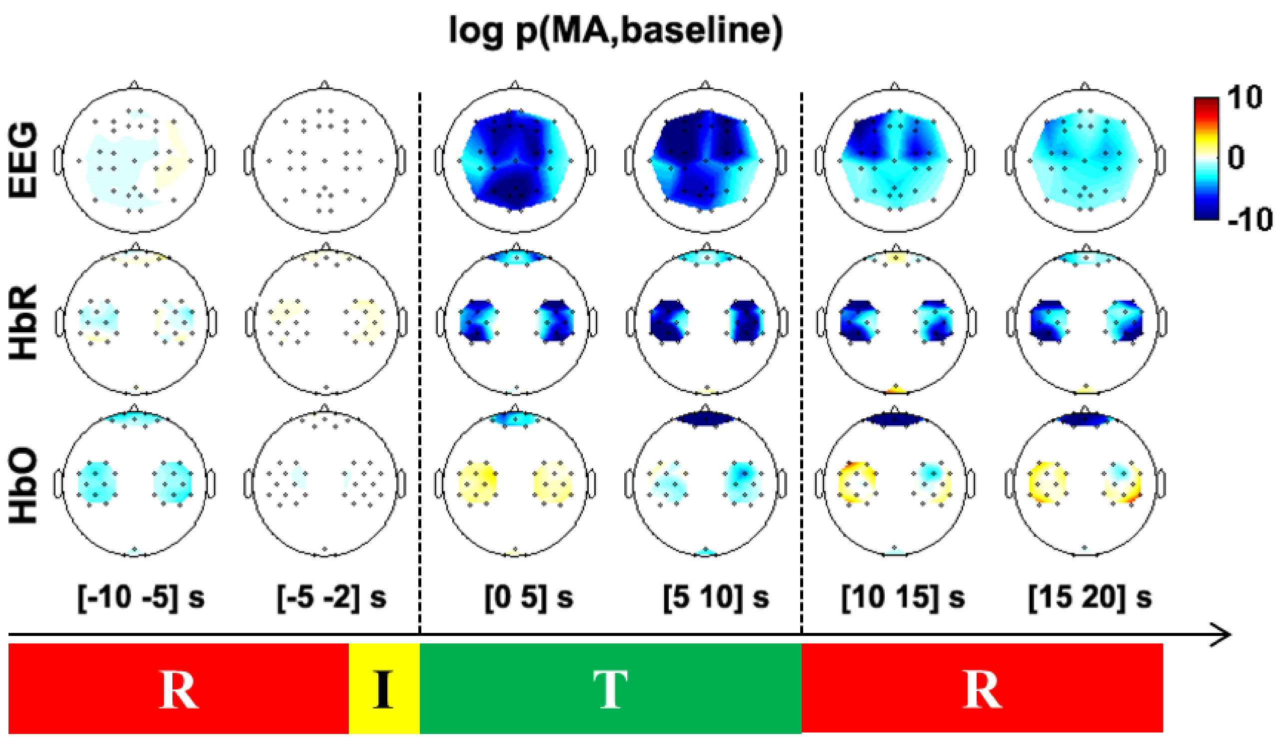

Fig. 5. Scalp evolution of grand average of log(p) based on signed r^2 for mental arithmetic in EEG and NIRS using the data of 15 subjects showing the best classification performance for better representation. Blue and Red colors denote higher separability of mental arithmetic and rest, respectively. The numbers displayed denote the time from the task onset. R, I and T denotes rest, instruction and task period, respectively. For better understanding, baseline correction was performed by subtracting the average value between -5 and -2 s.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2628057, IEEE Transactions on Neural Systems and Rehabilitation Engineering

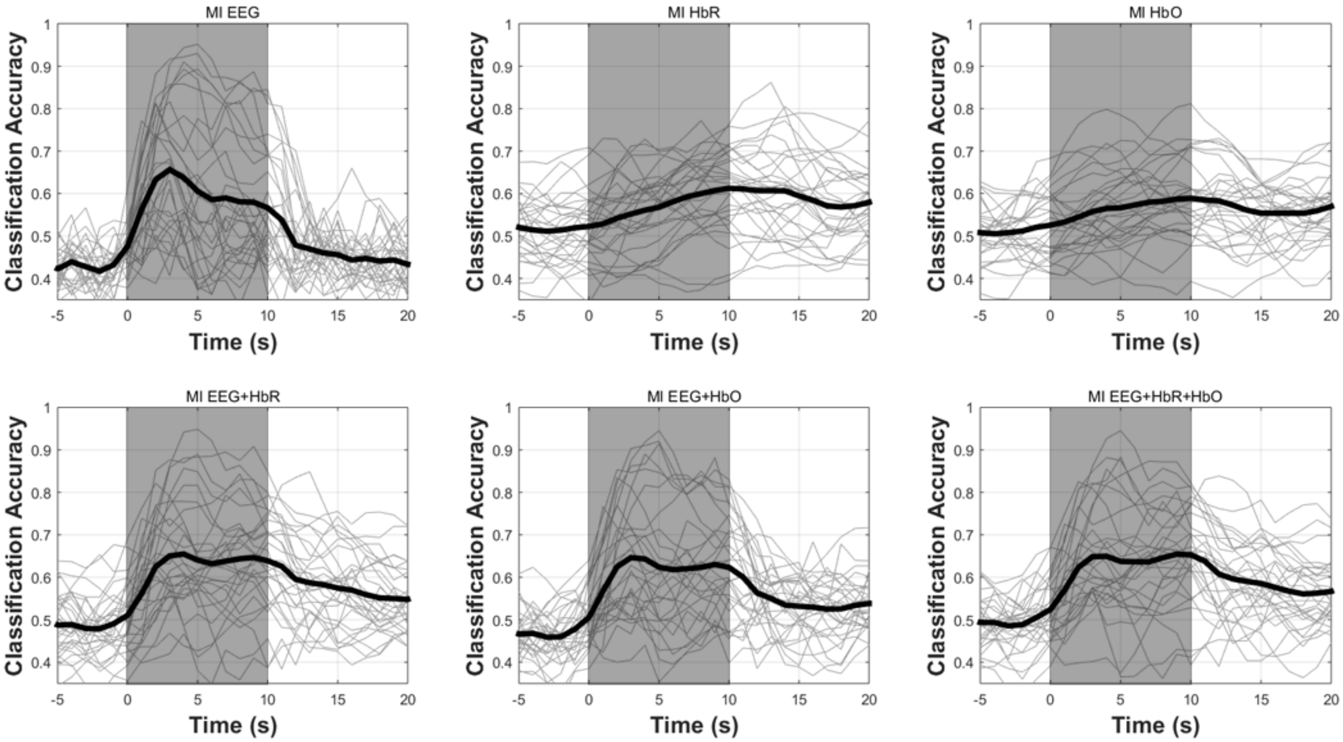

Fig. 6. EEG, NIRS and EEG+NIRS classification accuracies for 3 s moving time window for motor imagery (MI). The x-axis indicates the right edge of the moving time window. Gray lines show the individual classification accuracies while thick black lines show the average ones over whole subjects. The task starts at 0 s and finishes at 10 s. Gray shaded areas indicate task periods.

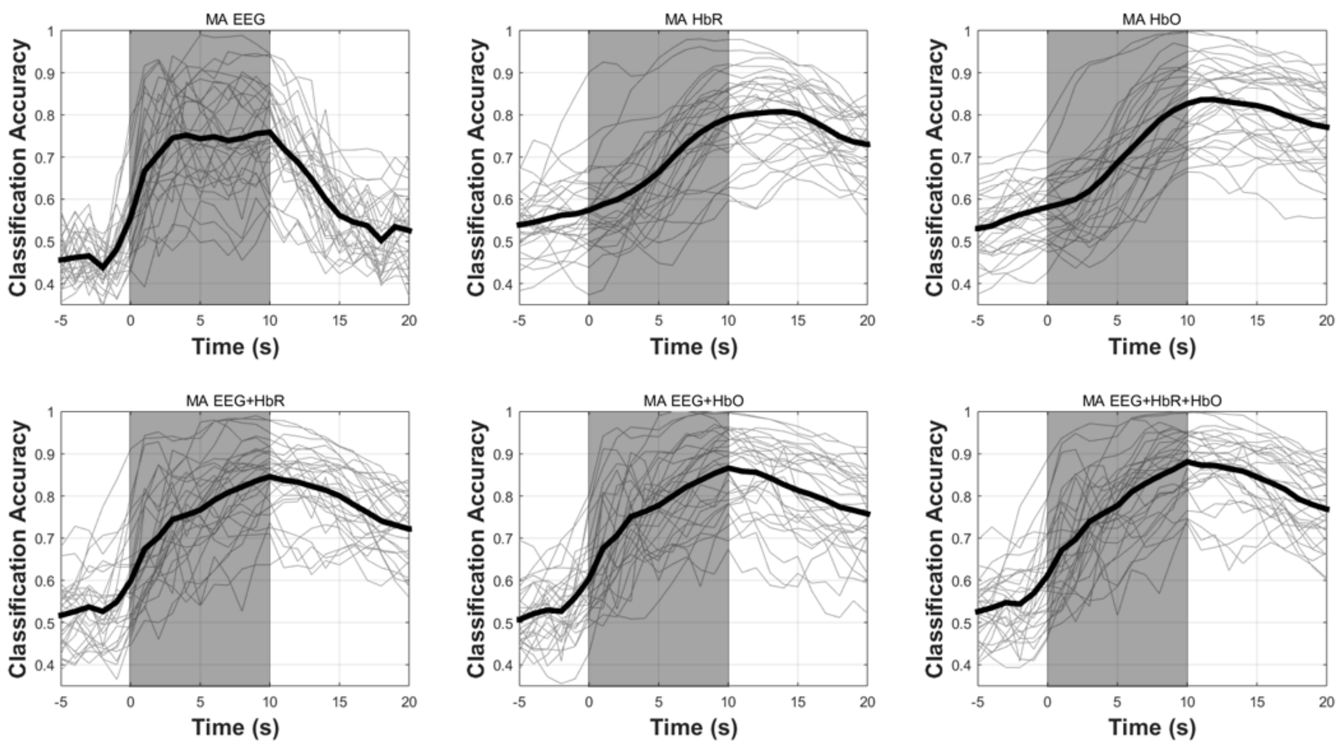

Fig. 7. EEG, NIRS and EEG+NIRS classification accuracies for 3 s moving time window for mental arithmetic (MA). The x-axis indicates the right edge of the moving time window. Gray lines show the individual classification accuracies while thick black lines show the average ones over whole subjects. The task starts at 0 s and finishes at 10 s. Gray shaded areas indicate task periods.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2628057, IEEE Transactions on Neural Systems and Rehabilitation Engineering

components containing the most discriminative information over each time period of the moving window. The moving time window was shifted from -5 s to 20 s based on the task onset. These features were used for constructing the feature vectors of each analysis time period. The dimension of each feature vector was four for discriminating left and right hand MI tasks (dataset A), and MA and baseline tasks (dataset B). For the NIRS data analysis, concentration changes of deoxy- and oxy-hemoglobin (HbR and HbO) were first calculated by the modified Lambert–Beer law, [34, 51]. The HbO and HbR data were band-pass filtered (6^{th} order zero-phase Butterworth filter with passband of 0.01 - 0.1 Hz). Baseline correction was performed by subtracting the average value between -5 and -2 s. All channels were used for further NIRS data processing. The average values of time courses of HbR and HbO and the average slope over the moving time window which is identical with the EEG analysis were used for creating feature vectors [52]. A shrinkage linear discriminant analysis (shrinkage LDA) was used as a classifier [49, 53]. A 10×5-fold cross-validation was performed to evaluate the classification performance for each analysis time window of the both data sets. Note that the CSP filters were built only based on training data, and then applied to test data of a classifier [17]. The same classifier and cross-validation approaches were applied to both EEG and NIRS data.

To figure out the potential merit of the hybrid BCI, EEG and NIRS-integrated data was evaluated based on a meta-classifier. After the estimation of the three individual classifiers (one for the EEG and two for the both chromophores of NIRS), we explored the classification performance of all possible combinations of EEG and both NIRS chromophores (i.e., EEG+HbR, EEG+HbO and EEG+both chromophores). The outputs of individual classifiers were combined to comprise feature vectors for the meta-classifier. We used the same shrinkage LDA as a meta-classifier. The cross-validation of the meta-classifier followed the same procedure as for the EEG and NIRS data analyses. For more details, please refer to [34].

F. Motion artifacts

No data processing was applied to EEG and NIRS (HbO) data measured during motion artifacts to show their original patterns in the time domain.

G. Statistical test

The statistical tests employed in this study are always performed using the ‘Wilcoxon signed rank sum test’.

III. Results

A. Grand average of log(p) significance based on signed squared point-wise biserial correlation coefficient

The scalp topographies of the grand averages of log(p) significance based on signed r^2 over 15 subjects (top 50% performers in decoding accuracy) are shown at specific time points for MI and MA in Figs. 4 and 5, respectively. A higher absolute value of log(p) means a better separability with higher significance between different conditions. During the MI task, EEG and HbR generally show good separability around the motor areas (Fig. 4). Note that the EEG is the best separable at a time interval of 0 - 5 s while HbR needs more time to reach the best separable time interval due to its inherent delayed responses (neurovascular coupling). HbO also shows significant separability around motor areas, but not as strong as EEG and HbR. This might be caused by unclassifiable MI-related subject-dependent activations. For the MA task in Fig. 5, EEG, HbR and HbO show high separability around frontal and parietal areas, and the delayed responses for NIRS is also observed in terms of separability. This result corresponds to a well-known neurophysiological behavior [54]. In HbR of Fig. 5, significant log(p) is observed particularly on motor area. As the rest interval of a preceding trial is not long enough to make the hemodynamic changes return to baseline state, HbR might be separable before the task onset at t = [-10 -5] s.

B. Classification accuracy: EEG, NIRS and Hybrid

Figs. 6 and 7 show EEG, NIRS and EEG+NIRS classification accuracies for the 3 s moving time window for MI and MA, respectively. The maximum values of the average EEG, HbR and HbO classification accuracies over time for MI are 65.6 %, 66.5 % and 63.5 %, respectively, while they reach to 75.9 %, 80.7 % and 83.6 % for MA. The MA-related activations were classifiable in more subjects than MI-related activations. For MI, only fourteen, nineteen and fifteen out of twenty-nine subjects achieved an acceptable classification accuracy for EEG, HbR and HbO, respectively, exceeding a theoretical chance level with statistical significance (67.5 %, $p = 0.05$) [55] while, for MA, twenty-five and twenty-nine subjects exceeded the chance level for EEG and both chromophores, respectively. Performance improvements in the form of a scatter plot are shown in Fig. 8(a) and (b). The unimodal EEG classification accuracies are plotted against those of EEG+HbR+HbO. The classification accuracies were selected at 10 s (end of task), a time point where the delayed NIRS signal is fully developed and the EEG signal is still active. More than 75.9 % and 86.2 % of the subjects showed improved performance by the hybrid approach for MI and MA tasks, respectively. The hybrid approach showed statistically significant improvement with respect to classification accuracy ($p < 0.001$). When comparing EEG classification accuracies with hybrid classification accuracies (EEG+HbR+HbO), the average performance using hybrid measures was increased in MI tasks by 8.6 % and MA tasks 12.2 %. The results underline that the hybrid approach in BCI is capable of enhancing the system performance.

C. Posterior questionnaire

Fig. 9 shows the average scores of the posterior questionnaire. Subjects reported that they could focus better on MA (MI: 3.24 ± 0.14, MA: 3.97 ± 0.15 (mean ± standard error), Wilcoxon signed rank sum test, $p < 0.001$). They got more tired (MI: 2.90 ± 0.18, MA: 2.45 ± 0.15, $p < 0.05$) and sleepy (MI: 2.93 ± 0.21, MA: 1.32 ± 0.16, $p < 0.001$) during MI than MA. They experienced more difficulty while performing MI, compared to MA (MI: 2.86 ± 0.21, MA: 2.21 ± 0.20, $p < 0.01$).

D. Motion artifact

Figs. 10(a) and (b) show the grand averages of EEG and NIRS (HbR and HbO) motion artifacts, respectively. The representative channels of each artifact type were selected by visual inspection. For NIRS, since the relevant BLK and MVE were not visible, the related figures were not included. The

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2628057, IEEE Transactions on Neural Systems and Rehabilitation Engineering

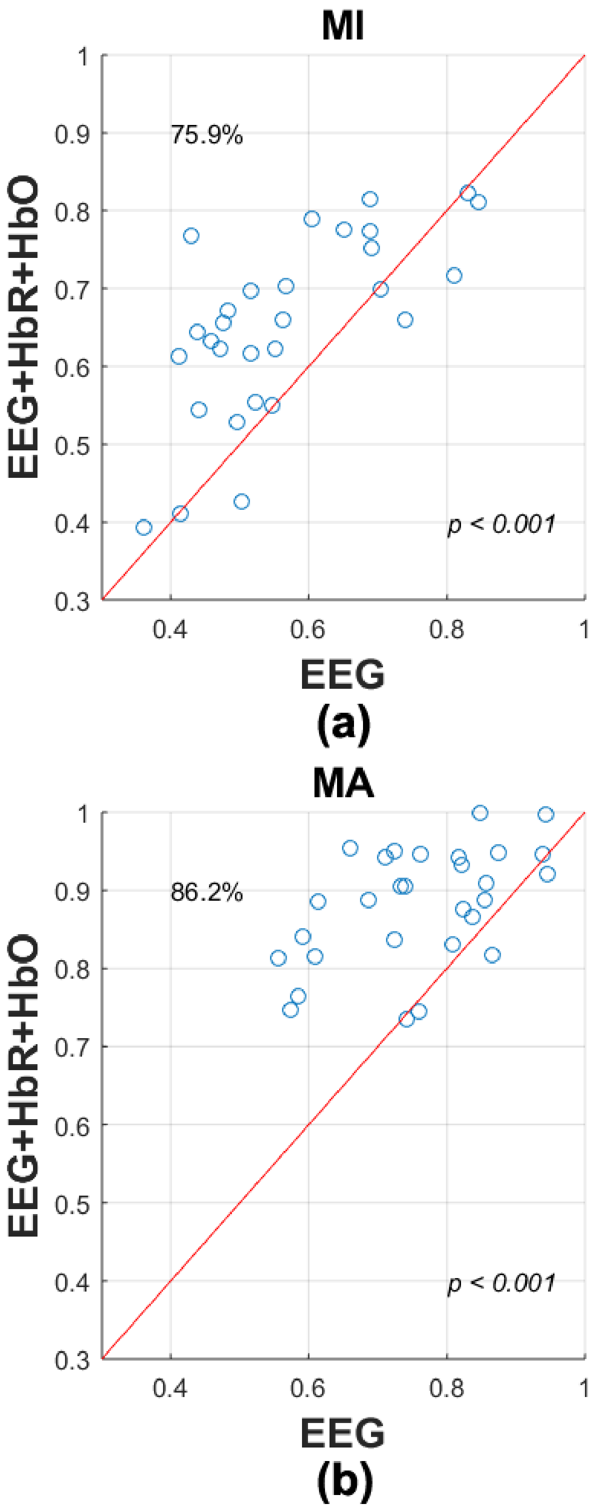

Fig. 8. Scatter plot of EEG classification accuracies (x-axis) against hybrid classification accuracies (y-axis) for (a) MI and (b) MA. Blue circles above the red diagonal line depict performance improvement by means of combinations of EEG and NIRS. The percentage values indicate the ratio of the number of subjects showing improved results to the whole subjects by means of the combinations. The *p*-values denote the significance of the improvement.

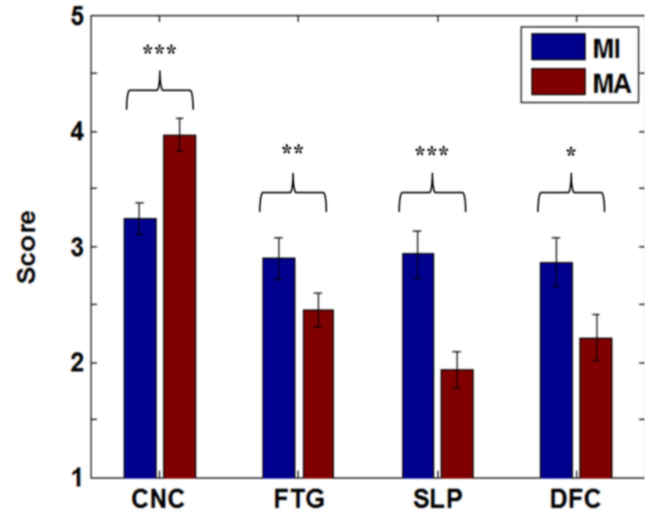

Fig. 9. Average scores of post-experiment survey regarding motor imagery (MI) and mental arithmetic (MA) with degree of significance ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$). CNC, FTG, SLP and DFC denote concentration, fatigue, sleepiness and difficulty, respectively.

EEG blink pattern averaged over all subjects is presented in the first panel of Fig. 10(a). Contrary to a typical blink pattern showing a large negative potential followed by its corresponding positive potential, an initial negative potential is relatively small. This is because the subjects blinked their eyes at a slightly different timing, thereby cancelling out an initial negative potential. However, the initial negative potential is clearly observed in the individual data (not shown here). When performing UD of MVE, position-specific vertical EOG is clearly detected at F7 showing three different levels for C, U1 and U2 positions (second panel in Fig. 10(a)). When performing UD of MVH, a similar pattern to UD of MVE is observed at F8 (third panel in Fig. 10(a)). Noisy oscillations are observed at T7 and T8 for CLT and OPM, respectively (fourth and fifth panels in Fig. 10(a)). For NIRS (Fig. 10(b)), HbO shows a larger motion artifact-related fluctuation than HbR, which indicates that HbO is more susceptible to motion artifacts than HbR [56].

IV. Discussion

A. Summary of the results

As indicated by the signed r^2 values, MI- and MA-related activations were classifiable over motor areas and fronto-parietal areas, respectively. For EEG, the classification rates rose without much delay, while for NIRS the classification accuracies reached the maximum after the delay of several seconds. MI-related activations of many subjects were not classifiable while classification of MA-related activations was better. During MA, the subjects could focus on the task more easily, which is presumably because the MA task was familiar to them. Also, classification performance significantly increased when simultaneously using EEG and NIRS signals, compared to that of each single modality.

Please note that this study did not aim to benchmark machine learning algorithms. It is very likely that the decoding accuracies achieved with our baseline signal processing approach could be outperformed by more sophisticated and state-of-the art methods, for instance those documented in [14, 15, 17, 43, 44]. Being aware that our approach led to rather poorer decoding accuracies, the results obtained from time segmented data can provide additional time-related information on the dataset. For a similar reason, we did not perform any channel or

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2628057, IEEE Transactions on Neural Systems and Rehabilitation Engineering

subject selection based on signal quality or decoding accuracy in the analysis, except for the illustration of log(p)-based scalp plots. However, to facilitate future approaches, we provide more details about our dataset such as statistics on the NIRS and EEG channel qualities and (individual) classification accuracies, also using the whole 10 s trials, in the supplementary material.

B. Particularity of the data set

As can be seen in Fig. 7, decoding of the mental state, especially MA, using NIRS signals is still well possible at 20 s after task onset. As the signals do not completely return to baseline, this makes decoding more challenging. Longer resting periods, for instance 28 s [57] up to 40 s [58] can prevent this phenomenon. However, to limit an extensive experimental runtime of > 3 h, we intentionally used a relatively short relaxation time. This should enable users of the dataset to tackle the well-known challenge in NIRS-based BCIs to achieve higher information transfer rate with tolerable trial lengths while using slow and typically up to 10 s delayed NIRS signals.

Moreover, the distribution of the task transitions is biased. Due to the way the task sequence was randomized, see Sec. 2.C.1, there are (on average) 25% of repetitions (a trial has the same task as the previous one) and 75 % of alternations. Therefore, a classifier trained in the pre-task interval could in principle identify the type of the previous trial from its carry-over effect and predict the opposite class for the subsequent trial, which would lead in the extreme case to a classification accuracy of 75%. In a weak way, we see this effect reflected in Fig. 7 for HbR and HbO where the classification accuracy is at about 55% already at -2 s. In the EEG, we see a different effect that might be surprising at first sight: a rise of classification accuracy from 50% already starting around -1 s relative to task onset. This is due to the fact that the task instruction was given 2 s before 'task onset', i.e., at -2 s. It is very likely that the subjects got ready to start as soon as they recognized the task, thereby producing detectable task-relevant activations before the task onset. Due to the same reason, some previous studies used data epoched prior to the task onset [59-61].

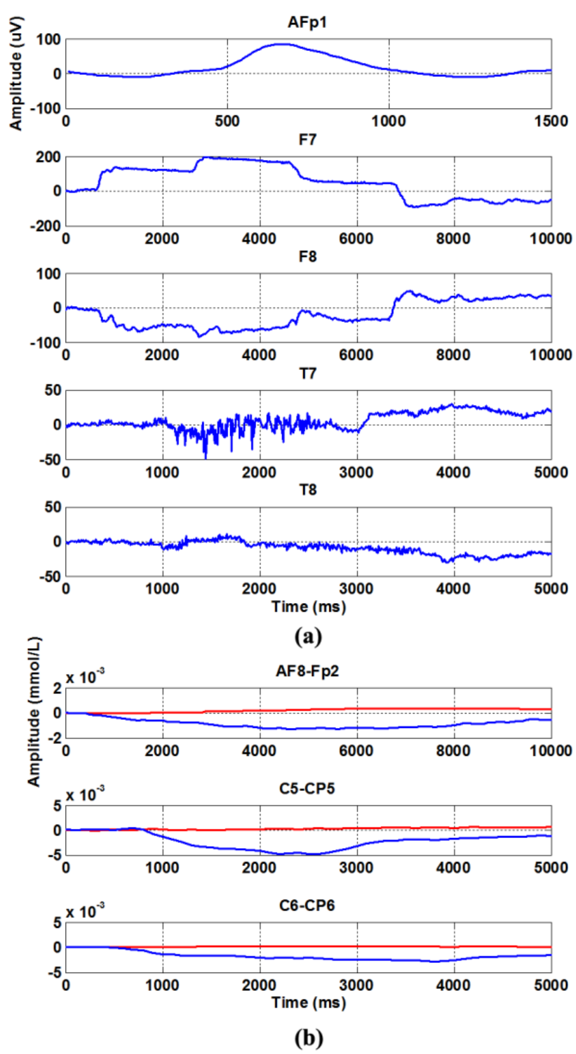

Fig. 10. Grand averages of (a) EEG and (b) NIRS motion artifacts (red: HbR, blue: HbO). For EEG, {BLK, MVE, MVH, CLT and OPM} are displayed from top to bottom. For NIRS, {MVH, CLT and OPM} are presented from top to bottom. EEG and NIRS channel names are shown on the top of each subplot. NIRS channel names are set as a concatenation of source and detector position. The near EEG and NIRS channels are selected to compare the signals for the same condition.

C. Benefit of hybrid BCI

Given that complementary information is measured by EEG and NIRS, their combination is capable of compensating the low performance of either and we can effectively improve the single-trial performance as presented in Fig. 7. In addition, in Fig. 6, HbR and HbO classification accuracies reached the threshold performance of a binary BCI (70 % accuracy [62, 63]) at 10 s and 8 s, respectively, while both EEG+HbR and EEG+HbO classification accuracies reached the threshold at 2 s. It is clear from these results that the hybrid BCI can compensate the inferior temporal responsiveness of NIRS and improve the performance in light of information transfer rate as well as classification accuracy as shown in [34].

D. Practical difficulty of hybrid BCI

Recent research reported that feedback training is helpful to enhance the performance of MI-based BCI [64]. However, subjects have to train for about one hour with a visual feedback indicating whether their performances work well or not. Also, a long preparation time is necessary to set up both EEG and NIRS systems, which is a main practical drawback to use MI-based hybrid BCIs. Therefore, it is difficult to finish feedback training and actual data recording in one day under time constraints and also difficult for subjects to maintain their concentration and good condition over the experiment. Even though the use of dry electrodes may reduce the EEG preparation time [65, 66], it is unavoidable to brush aside the subject's hair for the NIRS set-up. Also, signal quality is usually worse for dry electrodes than wet electrodes. Therefore, the development of a hybrid BCI based on prefrontal cortex signals could lead into a promising direction.

E. Advantages of an open access dataset for hybrid BCI

So far, because no EEG-NIRS hybrid dataset has been available to public, many researchers recorded and built new datasets to validate their proposed methods and algorithms. However, establishing a new dataset is usually time-consuming. Our open access dataset can save that time for acquiring a hybrid BCI dataset and thus help BCI researchers to more easily validate their novel BCI analysis methods.

V. Conclusion

We have provided baseline analysis results validating our open access dataset using standard signal processing methods as a reference guidance for a hybrid BCI. The dataset was established based on MATLAB as it is an analysis platform widely used in BCI research field. The signal processing was done using the BBCI toolbox [49] and EEGLAB [67]. We have managed the data as clearly as possible to facilitate it for research purposes. Both, the raw datasets are available for free download via: http://doc.ml.tu-berlin.de/hBCI.

Acknowledgement

We would like to thank Dr. Christoph Schmitz (NIRx GmbH, Berlin, Germany) for the fruitful advice and NIRx GmbH (Berlin, Germany) for providing a NIRS system. Also we would like to appreciate Dr. Jan Mehnert (Institute of Systems Neuroscience, Medical Center Hamburg-Eppendorf, Hamburg, Germany) for his helpful discussion.

Reference

[1] G. Dornhege, J. R. Millán, T. Hinterberger, D. McFarland, and K.-R. Müller, *Toward brain-computer interfacing*. Cambridge, MA: MIT press, 2007.
[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, Jun. 2002.
[3] J. R. Wolpaw and E. W. Wolpaw, *Brain-computer interfaces: principles and practice*. New York: Oxford University Press, 2012.
[4] X. R. Gao, D. F. Xu, M. Cheng, and S. K. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, pp. 137-140, Jun. 2003.
[5] T. Kaufmann, A. Herweg, and A. Kübler, "Toward brain-computer interface based wheelchair control utilizing tactually-evoked event-related potentials," *J. Neuroeng. Rehabil.*, vol. 11, p. 7, Jan. 2014.
[6] A. Ramos-Murguialday, D. Broetz, M. Rea, L. Laeer, O. Yilmaz, F. L. Brasil, *et al.*, "Brain-machine interface in chronic stroke rehabilitation: a controlled study," *Ann. Neurol.*, vol. 74, pp. 100-108, Jul. 2013.
[7] R. Ramirez and Z. Vamvakousis, "Detecting emotion from EEG signals using the emotive EPOC device," in *Brain Informatics*. vol. 7670, BI 2012 ed Berlin Heidelberg: Springer, 2012, pp. 175-184.
[8] P. Brunner, L. Bianchi, C. Guger, F. Cincotti, and G. Schalk, "Current trends in hardware and software for brain–computer interfaces (BCIs)," *J. Neural Eng.*, vol. 8, p. 025001, Apr. 2011.
[9] K.-R. Müller, M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, and B. Blankertz, "Machine learning for real-time single-trial EEG-analysis: from brain-computer interfacing to mental state monitoring," *J. Neurosci. Methods*, vol. 167, pp. 82-90, Jan. 2008.
[10] J. R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, *et al.*, "Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges," *Front. Neurosci.*, vol. 4, p. 00161, Sep. 2010.
[11] B. Blankertz, M. Tangermann, C. Vidaurre, S. Fazli, C. Sannelli, S. Haufe, *et al.*, "The Berlin brain–computer interface: non-medical uses of BCI technology," *Front. Neurosci.*, vol. 4, p. 00198, Dec. 2010.
[12] R. Krepki, B. Blankertz, G. Curio, and K.-R. Müller, "The Berlin brain-computer interface (BBCI) - towards a new communication channel for online control in gaming applications," *Multimed. Tools Appl.*, vol. 33, pp. 73-90, Apr. 2007.
[13] P. Sajda, E. Pohlmeyer, J. Wang, L. C. Parra, C. Christoforou, J. Dmochowski, *et al.*, "In a blink of an eye and a switch of a transistor: Cortically coupled computer vision," *Proc. IEEE*, vol. 98, pp. 462-478, Mar. 2010.
[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, pp. 41-56, Jan. 2008.
[15] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components—a tutorial," *Neuroimage*, vol. 56, pp. 814-825, May 2011.
[16] B. Blankertz, G. Dornhege, S. Lemm, M. Krauledat, G. Curio, and K.-R. Müller, "The Berlin brain-computer interface: Machine learning based detection of user specific brain states," *J. Univers. Comput. Sci.*, vol. 12, pp. 581-607, Jun. 2006.
[17] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller, "Introduction to machine learning for brain imaging," *Neuroimage*, vol. 56, pp. 387-399, May 2011.
[18] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 4, pp. R1-R13, Jun. 2007.
[19] *EEG open access datasets*. Available: http://bnci-horizon-2020.eu/database/data-sets
[20] P. Sajda, A. Gerson, K. R. Müller, B. Blankertz, and L. Parra, "A data analysis competition to evaluate machine learning algorithms for use in brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, pp. 184-185, Jun. 2003.
[21] B. Blankertz, K. R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, *et al.*, "The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1044-1051, Jun. 2004.
[22] B. Blankertz, K.-R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl, *et al.*, "The BCI competition III: validating alternative approaches to actual BCI problems," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, pp. 153-159, Jun. 2006.
[23] M. Tangermann, K.-R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, *et al.*, "Review of the BCI Competition IV," *Frontiers in Neuroscience*, vol. 6, p. 00055, Jul. 2012.
[24] B. D. Mensh, J. Werfel, and H. S. Seung, "BCI competition 2003-data set Ia: combining gamma-band power with slow cortical potentials to improve single-trial classification of electroencephalographic signals," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1052-1056, Jun. 2004.
[25] V. Bostanov, "BCI competition 2003-data sets Ib and IIb: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1057-1061, Jun. 2004.
[26] G. Blanchard and B. Blankertz, "BCI competition 2003-data set IIa: spatial patterns of self-controlled brain rhythm modulations," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1062-1066, Jun. 2004.
[27] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, "BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1073-1076, Jun. 2004.
[28] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang, "BCI competition 2003-data set IIb: enhancing P300 wave detection using ICA-based subspace projections for BCI applications," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1067-1072, Jun. 2004.
[29] S. Lemm, C. Schäfer, and G. Curio, "BCI competition 2003-data set III: probabilistic modeling of sensorimotor μ rhythms for classification of imaginary hand movements," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1077-1080, Jun. 2004.
[30] Y. Wang, Z. Zhang, Y. Li, X. Gao, S. Gao, and F. Yang, "BCI competition 2003-data set IV: an algorithm based on CSSD and FDA for classifying single-trial EEG," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1081-1086, Jun. 2004.
[31] G. Müller-Putz, R. Leeb, M. Tangermann, J. Höhne, A. Kübler, F. Cincotti, *et al.*, "Towards noninvasive hybrid brain–computer interfaces: framework, practice, clinical application, and beyond," *Proc. IEEE*, vol. 103, pp. 926-943, Jun. 2015.
[32] S. Fazli, S. Dähne, W. Samek, F. Bießmann, and K.-R. Müller, "Learning from more than one data source: data fusion techniques for sensorimotor rhythm-based brain–computer interfaces," *Proc. IEEE*, vol. 103, pp. 891-906, Jun. 2015.
[33] S. Dähne, F. Bießmann, W. Samek, S. Haufe, D. Goltz, C. Gundlach, *et al.*, "Multivariate machine learning methods for fusing multimodal functional neuroimaging data," *Proc. IEEE*, vol. 103, pp. 1507-1530, Sep. 2015.
[34] S. Fazli, J. Mehnert, J. Steinbrink, G. Curio, A. Villringer, K.-R. Müller, *et al.*, "Enhanced performance by a hybrid NIRS-EEG brain computer interface," *Neuroimage*, vol. 59, pp. 519-529, Jan. 2012.
[35] S. Amiri, R. Fazel-Rezai, and V. Asadpour, "A review of hybrid brain-computer interface systems," *Adv. Hum.-Comput. Interact.*, vol. 2013, p. 187024, Jan. 2013.
[36] A. P. Buccino, H. O. Keles, and A. Omurtag, "Hybrid EEG-fNIRS asynchronous brain-computer interface for multiple motor tasks," *Plos One*, vol. 11, p. 0146610, Jan. 2016.
[37] A. v. Lühmann, H. Wabnitz, T. Sander, and K.-R. Müller, "M3BA: a mobile, modular, multimodal biosignal acquisition architecture for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2016.2628057, IEEE Transactions on Neural Systems and Rehabilitation Engineering

miniaturized EEG-NIRS based hybrid BCI and monitoring," *IEEE Trans. Biomed. Eng.*, vol. PP, pp. 1-1, Sep. 2016.

[38] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proc. IEEE*, vol. 89, pp. 1123-1134, Jun. 2001.

[39] R. Sitaram, H. Zhang, C. Guan, M. Thulasidas, Y. Hoshi, A. Ishikawa, *et al.*, "Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface," *Neuroimage*, vol. 34, pp. 1416-1427, Feb. 2007.

[40] S. D. Power, T. H. Falk, and T. Chau, "Classification of prefrontal activity due to mental arithmetic and music imagery using hidden Markov models and frequency domain near-infrared spectroscopy," *J. Neural Eng.*, vol. 7, p. 026002, Apr. 2010.

[41] S. D. Power, A. Kushki, and T. Chau, "Automatic single-trial discrimination of mental arithmetic, mental singing and the no-control state from prefrontal activity: toward a three-state NIRS-BCI," *BMC Res. Notes*, vol. 5, p. 141, Mar. 2012.

[42] W. Samek, M. Kawanabe, and K.-R. Müller, "Divergence-based framework for common spatial patterns algorithms," *IEEE Rev. Biomed. Eng.*, vol. 7, pp. 50-72, Apr. 2014.

[43] F. Lotte and C. T. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms," *IEEE Trans. Biomed. Eng.*, vol. 58, pp. 355-362, Feb. 2011.

[44] N. A. Chadwick, D. A. McMeekin, and T. Tan, "Classifying eye and head movement artifacts in EEG signals," in *Proceedings of the 5th IEEE International conference on digital ecosystems and technologies conference (DEST)*, Daejeon, Korea, 2011, pp. 285-291.

[45] J. Dien, "Issues in the application of the average reference: Review, critiques, and recommendations," *Behav. Res. Methods Instr. Comput.*, vol. 30, pp. 34-43, Feb. 1998.

[46] G. Gomez-Herrero, W. De Clercq, H. Anwar, O. Kara, K. Egiazarian, S. Van Huffel, *et al.*, "Automatic removal of ocular artifacts in the EEG without an EOG reference channel," in *7th Nordic Signal Processing Symposium*, Reykjavik, Iceland, 2006, pp. 130-133.

[47] B. Blankertz, M. Kawanabe, R. Tomioka, F. Hohlefeld, K.-R. Müller, and V. V. Nikulin, "Invariant common spatial patterns: alleviating nonstationarities in brain-computer interfacing," in *Advances in neural information processing systems (NIPS)*, Vancouver and Whistler, Canada, 2007, pp. 113-120.

[48] Z. J. Koles and A. C. K. Soong, "EEG source localization: implementing the spatio-temporal decomposition approach," *Electroencephalogr. Clin. Neurophysiol.*, vol. 107, pp. 343-352, Nov. 1998.

[49] *BBCI toolbox*. Available: https://github.com/bbci/bbci_public/

[50] D. J. McFarland, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw, "Mu and beta rhythm topographies during motor imagery and actual movements," *Brain topogr.*, vol. 12, pp. 177-186, Mar. 2000.

[51] N. Kollias and W. Gratzer. (1999). *Tabulated molar extinction coefficient for hemoglobin in water*. Available: http://omlc.org/spectra/hemoglobin/summary.html

[52] N. Naseer and K.-S. Hong, "fNIRS-based brain-computer interfaces: a review," *Front. Hum. Neurosci.*, vol. 9, p. 00003, Jan. 2015.

[53] J. H. Friedman, "Regularized discriminant analysis," *J. Am. Stat. Assoc.*, vol. 84, pp. 165-175, Mar. 1989.

[54] G. Sammer, C. Blecker, H. Gebhardt, M. Bischoff, R. Stark, K. Morgen, *et al.*, "Relationship between regional hemodynamic activity and simultaneously recorded EEG-theta associated with mental arithmetic-induced workload," *Hum. brain mapp.*, vol. 28, pp. 793-803, Aug. 2007.

[55] E. Combrisson and K. Jerbi, "Exceeding chance level by chance: the caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy," *J. neurosci. methods*, vol. 250, pp. 126-136, Jul. 2015.

[56] X. Cui, S. Bray, and A. L. Reiss, "Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics," *Neuroimage*, vol. 49, pp. 3039-3046, Feb. 2010.

[57] G. Pfurtscheller, G. Bauernfeind, S. C. Wriessnegger, and C. Neuper, "Focal frontal (de) oxyhemoglobin responses during simple arithmetic," *Int. J. Psychophysiol.*, vol. 76, pp. 186-192, Jun. 2010.

[58] K. S. Hong, N. Naseer, and Y. H. Kim, "Classification of prefrontal and motor cortex signals for three-class fNIRS-BCI," *Neurosci. Lett.*, vol. 587, pp. 87-92, Feb. 2015.

[59] H.-J. Hwang, K. Kwon, and C.-H. Im, "Neurofeedback-based motor imagery training for brain–computer interface (BCI)," *J. Neurosci. Meth.*, vol. 179, pp. 150-156, Apr. 2009.

[60] N. F. Ince, S. Arica, and A. Tewfik, "Classification of single trial motor imagery EEG recordings with subject adapted non-dyadic arbitrary time-frequency tilings," *J. Neural Eng.*, vol. 3, pp. 235-244, Sep. 2006.

[61] B. Kamousi, A. N. Amini, and B. He, "Classification of motor imagery by means of cortical current density estimation and Von Neumann entropy," *J. Neural Eng.*, vol. 4, pp. 17-25, Jun. 2007.

[62] C. Vidaurre and B. Blankertz, "Towards a cure for BCI illiteracy," *Brain topogr.*, vol. 23, pp. 194-198, Jun. 2009.

[63] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, "Machine-learning-based coadaptive calibration for brain-computer interfaces," *Neural Comput.*, vol. 23, pp. 791-816, Mar. 2011.

[64] M. A. Lebedev and M. A. L. Nicolelis, "Brain−machine interfaces: past, present and future," *Trends Neurosci.*, vol. 29, pp. 536-546, Sep. 2006.

[65] C. Fonseca, J. S. Cunha, R. Martins, V. Ferreira, J. M. de Sá, M. Barbosa, *et al.*, "A novel dry active electrode for EEG recording," *IEEE Trans. Biomed. Eng.*, vol. 54, pp. 162-165, Jan. 2007.

[66] F. Popescu, S. Fazli, Y. Badower, B. Blankertz, and K.-R. Müller, "Single trial classification of motor imagination using 6 dry EEG electrodes," *Plos One*, vol. 2, p. 0000637, Jul. 2007.

[67] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. neurosci. meth.*, vol. 134, pp. 9-21, Mar. 2004.