

Bringing BCI into everyday life: Motor imagery in a pseudo realistic environment

Stephanie Brandl, Johannes Höhne, Klaus-Robert Müller, *Member, IEEE*, and Wojciech Samek, *Member, IEEE*

Abstract—Bringing Brain-Computer Interfaces (BCIs) into everyday life is a challenge because an out-of-lab environment implies the presence of variables that are largely beyond control of the user and the software application. This can severely corrupt signal quality as well as reliability of BCI control. Current BCI technology may fail in this application scenario because of the large amounts of noise, nonstationarity and movement artifacts. In this paper, we systematically investigate the performance of motor imagery BCI in a pseudo realistic environment. In our study 16 participants were asked to perform motor imagery tasks while dealing with different types of distractions such as vibratory stimulations or listening tasks. Our experiments demonstrate that standard BCI procedures are not robust to these additional sources of noise, implicating that methods which work well in a lab environment, may perform poorly in realistic application scenarios. We discuss several promising research directions to tackle this important problem.

I. INTRODUCTION

A Brain-Computer Interface (BCI) [1] provides a non-muscular communication system which detects the user's intents by measuring brain signals and directly converting them into control signals. Several diseases, such as amyotrophic lateral sclerosis (ALS), lead to loss of muscular control. In late stage of such diseases, people are affected by complete *locked-in* syndrome, which means they have lost any voluntary muscular control and are therefore unable to communicate. Their completely functional minds are captured in paralyzed bodies. A BCI aims to give such patients new opportunities to communicate.

Various BCI studies have been carried out over the last few years [2] [3] [4]. By integrating machine learning algorithms, the duration of calibration time could be reduced significantly [5], resembling a highly important step towards usability

S. Brandl (s.brandl@mailbox.tu-berlin.de), is with the Berlin Institute of Technology, Marchstr. 23, 10587 Berlin, Germany.

J. Höhne (j.hoehne@tu-berlin.de) is with the Berlin Institute of Technology, Marchstr. 23, 10587 Berlin, Germany.

K.-R. Müller (klaus-robert.mueller@tu-berlin.de) is with the Berlin Institute of Technology, Marchstr. 23, 10587 Berlin, Germany and with the Department of Brain and Cognitive Engineering, Korea University, Anam-dong, Seongbuk-gu, Seoul 136-713, Korea.

W. Samek (wojciech.samek@hhi.fraunhofer.de) is with the Fraunhofer HHI, Einsteinufer 37, 10587 Berlin, Germany.

*This work was supported by the Federal Ministry of Education and Research (BMBF) under the project Adaptive BCI (FKZ 01GQ1115), by the DFG, by the World Class University Program through the National Research Foundation of Korea funded by the Ministry of Education, Science, and Technology, under Grant R31-10008 and by the Brain Korea 21 Plus Program through the National Research Foundation of Korea funded by the Ministry of Education.

of BCIs. BCI efficiency could be enhanced, especially for spelling devices, which now work with an accuracy of up to 99% [3]. Also, motor imagery-based BCIs have shown great improvements due to novel approaches in terms of robust feature extraction [6], robust artifact classification [7] and adaptive [8] methods, which have also been applied to locked-in patients [9].

However, most studies usually take place in situations where participants are sitting still and can focus completely on the respective task. These very controlled environments often do not reflect everyday life situations, where people also have to deal with many kinds of visual, auditory or other distractions. Since the main purpose of a BCI is to offer disabled people a new way of communication, it is highly necessary to start bringing the BCI research out of these controlled environments and into the real world. Therefore, we conducted a systematical artifact study for BCI.

In our study, participants not only needed to perform a motor imagery task but also had to deal with different kinds of distractions to simulate a pseudo-realistic environment. They were asked to close their eyes, listen to news, search the room for a particular number, watch a flickering video or deal with vibro-tactile stimulation. The purpose of this study was to evaluate the performance of standard BCI techniques in this pseudo-realistic environment. We also investigated the effects of the different types of distractors on motor imagery performance and, additionally, showed that it is in some cases possible to classify them against each other.

II. EXPERIMENTS

A. Participants

Sixteen healthy volunteers (6 female; range: 22-30 years; mean age: 26.3 years) participated in the study. Three of them already had BCI experience, though only one also had participated in a motor imagery experiment. All participants spoke and understood German at a level which allowed them to easily follow the German auditory and visual instructions. Three members of the TU Berlin Machine Learning Group voluntarily took part in the study, whereas the other volunteers were paid for their participation.

B. Distractions

Five different distractions and a control condition without any additional noise were included in the experimental setup to distract the person performing motor imagery tasks. In

the following we explain the design and motivation for these distractions.

Clean: No distraction was added, making this condition the control group.

Closed Eyes: Participants were asked to close their eyes during the motor imagery task. The main intention behind this was to see the effect of a more prominent alpha rhythm, which, after spectral filtering, may overlay the motor task-related mu rhythm, since both appear within a similar frequency band (8-13Hz). This condition was the main reason for giving the instructions and feedback over headphones instead of visually.

News: A news sequence was played during four trials over the headphones. For this purpose, parts of a public German newscast (Tagesschau) had been cut into sequences of 26 seconds, containing both, current news and news from 1994. Every news sequence was played exactly once (except for participant *od*, for whom some files were played twice) in each experiment. The purpose of this task was to distract the participant and to also investigate the influence of the activated auditory cortex on motor imagery performance. The volume was adapted such that it was still possible to clearly receive the instructions.

Numbers: A window appeared on the screen requesting the participant to find a certain letter to match with a stated number. For this purpose, 26 pieces of paper had been put up on the wall in front of the participant as well as on the right and left side of the room, so that it was necessary to turn the head in order to see the different sheets of paper. The numbers from 1 to 26 and all letters from the classical Latin alphabet were randomly mixed together in a way that each piece of paper had a combination of one capital letter and a number written on it. The participant was asked to search the room for the respective sheet with the correct combination of number and letter. If the participant found it, he was expected to say the letter out loud so we could count the found letters. Each combination was connected to one trial, such that the window changed after 7 seconds and the participant then had to search for another letter. Each combination was shown two or three times during one experiment. Although we counted the found letters, the main task was still to imagine the respective hand movement. The average across all 16 experiments of found letters was 59.7 out of 72, in a range from 40 to 68. We included this task to investigate the effect of additional muscular artifacts and additional distraction on the results of the motor imagery task.

Flicker: A video was played on a screen with a flicker in gray shades which altered the background at a frequency of 10 Hz. The purpose of this task was to identify the influence of the *steady state visually evoked potential* (SSVEP) [10] on the classification rate.

Stimulation: On the insides of both forearms of each participant, two vibration tactiles with a diameter of 3cm were taped, one approximately over the wrist and the other just below the elbow, such that it was still possible to comfortably bend the arm. Vibratory stimulation was carried out using

carrier frequencies of 50 and 100Hz, each modulated at 9, 10 and 11Hz respectively. The main goal here was to analyze the influence of the *steady state vibration somatosensory evoked potential* (SSVSEP) which has already been investigated in [11] and used for BCI in [12]. We expected an overlap in the somatosensory area of the evoked potential, due to vibration stimulation, and ERS/ERD effects, due to motor imagery.

C. Experimental Setup

During the experiment, the participant sat in an armchair at a distance of about 1m away from a 24" (16:10) computer screen. Furthermore, the participant was wearing headphones to receive auditory instructions.

A Fast'n Easy Cap (EasyCap GmbH) with 63 wet Ag/AgCl electrodes placed at symmetrical positions based on the International 10-20 system [13] was used for recording the EEG signals. The signals were referenced to the nose. Two 32-channel amplifiers (Brain Products) were used to amplify the signals, which were sampled at 1000 Hz.

The recording lasted about 90 minutes. Including the time needed for preparing the cap and the breaks in between, each experimental session lasted about three hours. Before the main experiment started, some baseline EEG recordings were carried out, during which the participant had to alternately open and close both eyes for about 15 seconds, with 4 repetitions each.

The main experiment was divided into 7 runs, each run lasting about 10 minutes. The first run was used as a calibration phase, where no feedback was given and no distractions were added. Each run after the calibration phase included 12 trials per distraction (including the *clean* condition) with each trial lasting about 4.5 seconds and consisting of one motor imagery task. The participant received instructions over the headphones at the beginning of each trial, which consisted of *left* and *right* announcements (since they were in German, they actually were *links* and *rechts* announcements). After 4.5 seconds, there was a *stop* announcement followed by a break, which lasted 2.5 seconds, before the next trial started. Every three to four minutes there was a possibility for a larger break if desired.

During the calibration phase, Laplacian filters [14] of electrodes C3 and C4 were calculated and an LDA classifier (linear discriminant analysis) was trained upon the spectral power of the signals in a broad band (9-13Hz and 18-26 Hz). Since the intention was to compare different methods, none of them was used during the experiment so that no participant could adapt a strategy to the respective method. The main purpose for the online classification was to give the volunteers some kind of feedback, hoping that this would keep their motivation levels high.

After the calibration phase, the trained classifier was applied to classify the imagined hand movements and to provide auditory feedback through the classifier's decision. This means that the trial started as before with a *left* or *right* instruction, and was ended by the *stop* announcement, followed by *decision left* (*Entscheidung links*) or *decision right* (*Entscheidung rechts*) during the 2.5 seconds break.

TABLE I: Mean classification accuracies for all distractions. For each experiment, the conditions with highest (bold) and lowest (red) performance rates are highlighted.

csp	od	njy	njz	nkm	nko	nkq	nkt	obx
overall	93.06	64.58	74.77	57.87	74.25	60.55	68.75	70.14
clean	95.83	76.39	88.89	63.89	72.22	63.89	72.22	80.56
eyesclosed	95.83	61.11	81.94	52.78	69.44	59.72	76.39	68.06
news	94.44	66.67	80.56	56.94	79.17	59.22	70.83	73.61
numbers	84.72	48.61	52.78	52.78	76.39	50.00	62.50	62.50
flicker	97.22	69.44	79.17	61.11	81.94	72.22	76.39	72.22
stimulation	90.28	65.28	65.28	59.72	66.16	58.33	54.17	63.89

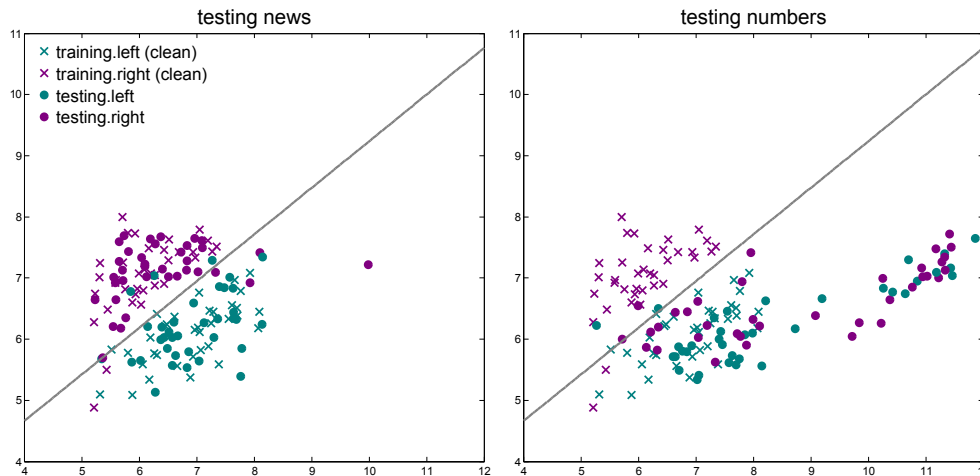


Fig. 1: Features (participant *njz*) of the classifications between left and right hand motor imagery (two best CSP filters) where CSP was only trained on *clean* and tested on *news* or *numbers*, respectively.

The motor imagery task during the online phase remained the same, except for the different distractions which were integrated into the trials.

D. Data Analysis

For the offline analysis, the data was downsampled to 100Hz and lowpass filtered in the range of 8-30Hz. We then cut the data into epochs, from 750ms to 3500ms, referring to the onset of each trial. We used Common Spatial Patterns (CSP) [15] to extract the features (three filters per class) and to train a classifier based on LDA. Since most of the participants were using the BCI for the first time, and some tasks were rather demanding, not everyone achieved classification rates significantly higher than chance level. By applying a binomial test ($\alpha = 0.05$), we determined a threshold of 61.11% over which we could assume actual BCI control. We selected the 8 experiments with classification rates higher than that threshold in the control group (*clean*) for further analysis.

III. EVALUATION

Classification results for the different conditions are summarized in Table I. Although all displayed participants achieved significant classification rates when no distractions were added, there were still several classification rates lower

than threshold. As expected, the *numbers* task, where users not only had to concentrate on a certain number but also had to turn their head to find the number in the room, was the one where participants had the most difficulties. Meanwhile, in the *news* or the *flicker* task, almost everyone achieved a performance rate higher or equal to their overall mean. So neither the user’s visual or auditory distraction, nor the changed brain signals, due to SSVEP effects or an activated auditory cortex, respectively, seem to have had a major impact on performance rates. Participant *njz* even achieved an almost 30% higher classification rate in the *news* and *flicker* tasks than in the *numbers* task. To investigate those large differences, we plotted training and testing features (see Fig. 1), with training features only consisting of *clean* motor imagery and testing only of *news* or *numbers* motor imagery. While *news* features show similar distribution to the *clean* training features, *numbers* features demonstrate significant shifts, which are likely the reason for the resulting low classification rate.

We also classified the different distractions against each other (one CSP filter per condition), separately for both hands. To get a better understanding of the distribution shifts compared to motor imagery without distractions, we visualized the mean classification rates over all 8 experiments

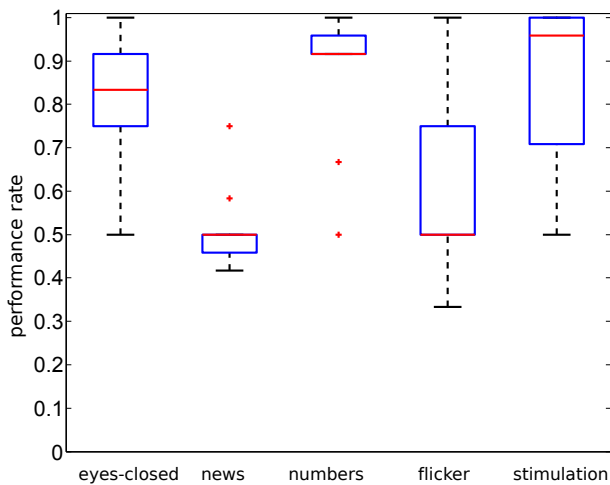


Fig. 2: Mean classification accuracies across all 8 experiments under different distraction conditions against *clean* motor imagery.

of the different tasks against left hand *clean* motor imagery, as shown in Fig. 2. As expected, feature shifts in the *numbers* task yield very high classification rates against *clean* motor imagery, whereas the expected similarities between *clean* and *news* motor imagery resulted in a classification rate at around chance level.

IV. CONCLUSION

Lab environments are highly artificial and systematically deprived of the sensory richness encountered in natural environments. Although, due to nearly perfect control over environmental factors, they offer a great testbed for the development and evaluation of algorithms, it needs to be kept in mind that research findings do not easily translate to real-world scenarios. The study presented in this paper has investigated the gap between the experimental conditions found in a lab and the conditions found in everyday life environments.

Our analysis clearly demonstrates that standard BCI procedures cannot handle the additional noise sources potentially encountered in out-of-lab settings as different types of distraction may greatly affect extracted features and lead to nonstationarity. This induced nonstationarity is the main reason for drops in classification accuracy.

Although initial attempts have been made to improve BCI performance in out-of-lab environments, more systematic studies are needed. Promising research directions seem to be robust and invariant BCI approaches [16], [17], [18], multi-subject methods [19], [20], [21] and adaptive algorithms [22].

V. ACKNOWLEDGEMENTS

We thank Benjamin Blankertz for valuable discussions on the experimental design, Claudia Sannelli for sharing her expertise on motor imagery based BCIs and Alexander von Lühmann for useful comments on the manuscript. Also, we thank all the participants for taking part in the study.

REFERENCES

- [1] G. Dornhege, J. del R. Millán, T. Hinterberger, D. McFarland, and K.-R. Müller, Eds., *Toward Brain-Computer Interfacing*. Cambridge, MA: MIT Press, 2007.
- [2] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, "A spelling device for the paralysed," *Nature*, vol. 398, no. 6725, pp. 297–298, 1999.
- [3] M. Thulasidas and C. Guan, "Optimization of bci speller based on p300 potential," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the IEEE*, 2005, pp. 5396–5399.
- [4] L. Acqualagna and B. Blankertz, "Gaze-independent bci-spelling using rapid serial visual presentation (rsvp)," *Clinical Neurophysiology*, vol. 124, no. 5, pp. 901–908, 2013.
- [5] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive berlin brain-computer interface: fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539–550, 2007.
- [6] W. Samek, M. Kawanabe, and K.-R. Müller, "Divergence-based framework for common spatial patterns algorithms," *IEEE Reviews in Biomedical Engineering*, vol. 7, pp. 50–72, 2014.
- [7] I. Winkler, S. Brandl, F. Horn, E. Waldburger, C. Allefeld, and M. Tangermann, "Robust artifactual independent component classification for bci practitioners," *Journal of neural engineering*, vol. 11, no. 3, p. 035013, 2014.
- [8] P. Shenoy, M. Krauledat, B. Blankertz, R. P. Rao, and K.-R. Müller, "Towards adaptive classification for bci," *Journal of neural engineering*, vol. 3, no. 1, p. R13, 2006.
- [9] J. Höhne, E. Holz, P. Staiger-Sälzer, K.-R. Müller, A. Kübler, and M. Tangermann, "Motor imagery for severely motor-impaired patients: evidence for brain-computer interfacing as superior control solution," *PLoS one*, vol. 9, no. 8, p. e104854, 2014.
- [10] J. Ding, G. Sperling, and R. Srinivasan, "Attentional modulation of ssvp power depends on the network tagged by the flicker frequency," *Cerebral cortex*, vol. 16, no. 7, pp. 1016–1029, 2006.
- [11] S. Tobimatsu, Y. M. Zhang, and M. Kato, "Steady-state vibration somatosensory evoked potentials: physiological characteristics and tuning function," *Clinical neurophysiology*, vol. 110, no. 11, pp. 1953–1958, 1999.
- [12] A.-M. Brouwer and J. B. Van Erp, "A tactile p300 brain-computer interface," *Frontiers in neuroscience*, vol. 4, p. 19, 2010.
- [13] H. Jasper, "The ten twenty electrode system of the international federation," *EEG Clin. Neurophysiol.*, vol. 10, pp. 371–375, 1958.
- [14] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for eeg-based communication," *Electroencephalography and clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997.
- [15] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial eeg during imagined hand movement," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 4, pp. 441–446, 1998.
- [16] B. Blankertz, M. Kawanabe, R. Tomioka, F. U. Hohlefeld, V. Nikulin, and K.-R. Müller, "Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing," in *Advances in Neural Information Processing Systems 20 (NIPS)*, 2008, pp. 113–120.
- [17] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 355 – 362, 2011.
- [18] W. Samek, C. Vidaurre, K.-R. Müller, and M. Kawanabe, "Stationary common spatial patterns for brain-computer interfacing," *Journal of Neural Engineering*, vol. 9, no. 2, p. 026013, 2012.
- [19] F. Lotte and C. Guan, "Learning from other subjects helps reducing brain-computer interface calibration time," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2010, pp. 614–617.
- [20] M. Krauledat, M. Tangermann, B. Blankertz, and K.-R. Müller, "Towards zero training for brain-computer interfacing," *PLoS ONE*, vol. 3, no. 8, p. e2967, 2008.
- [21] W. Samek, F. C. Meinecke, and K.-R. Müller, "Transferring subspaces between subjects in brain-computer interfacing," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, pp. 2289–2298, 2013.
- [22] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, "Machine-learning based co-adaptive calibration," *Neural computation*, vol. 23, no. 3, pp. 791–816, 2011.