Online Detection of P300 related Target Recognition Processes During a Demanding Teleoperation Task

Classifier Transfer for the Detection of Missed Targets

Hendrik Woehrle[1] and Elsa Andrea Kirchner[1,2]

[1] *Robotics Innovation Center, German Research Center for Artificial Intelligence (DFKI GmbH) Robert-Hooke-Str 5, Bremen, Germany*

[2] *Robotics Lab, University of Bremen, Robert Hooke Str. 5, Bremen, Germany*

Keywords: Brain Computer Interfaces, Embedded Brain Reading, P300, Single Trial Detection, Exoskeleton.

Abstract: The detection of event related potentials and their usage for innovative tasks became a mature research topic in the last couple of years for brain computer interfaces. However, the typical experimental setups are usually highly controlled and designed to actively evoke specific brain activity like the P300 event related potential. In this paper, we show that the detection and passive usage of the P300 related brain activity is possible in highly uncontrolled and noisy application scenarios where the subjects are performing demanding senso-motor task, i.e., telemanipulation of a real robotic arm. In the application scenario, the subject wears an exoskeleton to control a robotic arm, which is presented to him in a virtual scenario. While performing the telemanipulation task he has to respond to important messages. By online analysis of the subject's electroencephalogram we detect P300 related target recognition processes to infer on upcoming response behavior or missing of response behavior in case a target was not recognized. We show that a classifier that is trained to distinguish between brain activity evoked by recognized task relevant stimuli and ignored frequent task irrelevant stimuli can be applied to classify between brain activity evoked by recognized task relevant stimuli and brain activity that is evoked in case that task relevant stimuli are *not* recognized.

1 INTRODUCTION

Online-analysis and detection of specific patterns in electroencephalographic (EEG) data has been used for various applications, e.g., brain computer interfaces (BCIs). Current EEG-based BCIs are using classification and data dependent signal processing methods to detect the patterns in the EEG. Therefore, they highly depend on *training data* that has to be used for the calibration of the system before they can be used to detect the patterns in the *application data*. Usually, the training data has to be *subject-specific*, i.e., it has to be acquired from the subject in training sessions directly before the usage of the system. Further, the recorded data should be *clean*, i.e., free of artifacts that might affect the training or detection process, as well as *task specific*, i.e., must consist of data that is directly related to the patterns that are supposed to be detected. Therefore, most results are conducted in highly controlled artificial scenarios, where most of the possible disturbance sources have been excluded by experimental design and the subject may even be fixed in a specific position.

For many applications this is no draw back, especially if BCIs are applied as active interfaces, i.e., to control a machine or computer (Farwell and Donchin, 1988; Guger et al., 1999; Wolpaw et al., 2002; Reuderink, 2008; Nijholt et al., 2008). If, however, the patterns that should be detected in the brain activity are no longer actively produced, as it is the case for passive BCIs (Zander et al., 2010; George and Lécuyer, 2010), then background EEG that is evoked by the active task may overlay with the relevant brain. Since the subject is possibly performing different tasks, the background EEG may differ strongly depending on the situation and performed active task or action of the user and thus affects the training data.

For future practical applications of passive approaches (Zander et al., 2010; George and Lécuyer, 2010; Kirchner et al., 2010; Haufe et al., 2011; Kirchner and Drechsler, 2013), it is required to expose the systems and subjects to concrete, realistic use-cases, that are more uncontrolled and performed in perturbed environments. These conditions likely in-

Woehrle H. and Kirchner E..
Online Detection of P300 related Target Recognition Processes During a Demanding Teleoperation Task - Classifier Transfer for the Detection of Missed Targets.
DOI: 10.5220/0004724600130019
In *Proceedings of the International Conference on Physiological Computing Systems* (PhyCS-2014), pages 13-19
ISBN: 978-989-758-006-2

crease the amount of noise in the training as well as application data and may therefore impair the detection accuracy.

A further problem exists, if the amount of training data is small. It might not be possible to acquire a large amount of training examples in complex application scenarios. This is a general problem, since the detection accuracy of data-dependent signal processing and classification methods depends on the amount of available training data. Hence, approaches that can handle a reduced amount of training examples must be developed and applied.

One approach is to transfer a classifier between classes, i.e., to perform a classifier transfer (Pan and Yang, 2010). It could be shown that classifier transfer for the detection of patterns in the EEG performs well for a transfer of classifier between tasks in which the same event related activity has to be detected (Iturrate et al., 2013) or between similar types of event related potentials (ERPs) like different types of error potentials (Kim and Kirchner, 2013). In a recent work we showed that the transfer of classifier is also possible between classes that "miss" a pronounced pattern, i.e., the P300 (Kirchner et al., 2013). Hence, the data processing methods (classifiers and spatial filters) need not to be trained and tested on examples that are evoked by the same brain processes, like same or similar error detection processes, but by brain processes that evoke brain pattern, which are similar in shape and characteristics, i.e., miss a prominent ERP or pattern of ERPs.

By now, our investigations have been conducted in controlled experimental setups in an offline fashion. In this paper, we investigate the ability to detect the P300 ERP in a demanding dual task application scenario that combines an oddball paradigm with a second task. We show that the detection of P300 related target recognition processes and even more important the *missing* of target recognition processes can be performed online while a subject is performing a demanding and realistic interaction task that occupies the operators attention. This task consists of the teleoperation of a real robotic arm through a labyrinth via a virtual immersion scenario.

The paper makes the following contributions: 1) we demonstrate that the online, single trial detection of the P300 potential is possible in an application scenario that is affected by a high number of noise sources and artifacts and requires dual task performance from the subject (Kirchner and Kim, 2012), i.e., distracts the subject from the perception of task relevant stimuli; 2) we show that the few number of examples of training data of a specific class can be compensated to a certain degree by classifier transfer.

2 APPLICATION SCENARIO

In the proposed application scenario, we investigate whether it is possible to reliably detect target recognition processes as well as the missing of target recognition processes while a subject is performing a demanding teleoperation task.

Precisely, the experimental setup was as follows (see Fig. 1): The subjects were wearing an exoskeleton that covered their back and right arm (Folgheraiter et al., 2012), and a smart glove on their hand that were used as input devices for the teleoperation task.

In addition, participants were equipped with a head mounted display (HMD) on which the teleoperation site (including surroundings, labyrinth and robot) could be seen in 3D. Additionally to the 3D environment, information from the control system, a camera picture of the real scene and tools like a gyroscope depicting the orientation of the end-effector were at any time in the operators field of view. Head and hand movements of the operator were tracked (InterSense, Billerica, USA) and used to update the HMD.

The subjects had two main tasks that had to be performed at the same time: a) to control a robotic arm (*teleoperation task*) using the exoskeleton, and b) to respond to specific messages (*oddball task*).

2.1 The Teleoperation Task

In the teleoperation task, the end-effector of a robotic arm had to be steered through a labyrinth (see Fig. 1 C). This task is similar to a wire loop game, i.e., a certain path has to be followed and touching the labyrinth had to be avoided. The movements of the robotic arm were controlled via the exoskeleton by mapping the state and relative position of the exoskeleton components to a Mitubishi PA-10 robotic arm (see Fig. 1 A in the lower right corner) via a virtual model (see Fig. 1 A in the upper right corner) thereof (depending on the concrete type of investigation, see Sec. 3).

The teleoperation task is difficult and demanding for the subject, and therefore forces the subject to concentrate on it. Further, the subject was requested to rest from time to time. In each run 24 ± 8 rest periods had to be performed (Seeland et al., 2013). During rest the active exoskeleton kept the operators arm in position. While this was the case the operator was not allowed to respond to any warning (infrequent task relevant stimuli, see below) that were presented to him in an oddball fashion throughout the run.

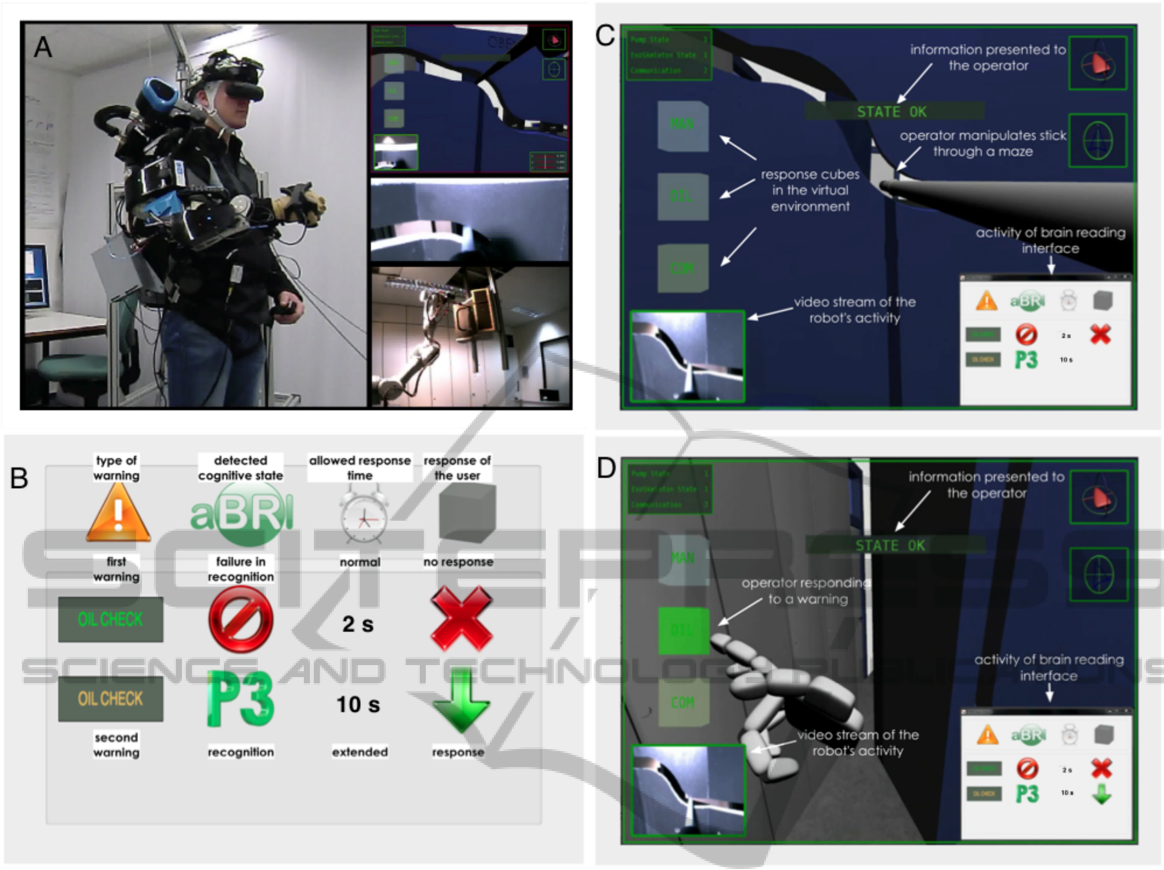

Figure 1: **Teleoperation scenario:** (A) An operator controls a robot via an exoskeleton and a virtual environment; (B) The operator monitoring system (OMS) integrates all important informations, it supervises brain states detected by brain reading, monitors the events in the scenario and responses of the operator on target events; the allowed response time is adapted with respect to brain state, i.e., whether a warning was recognized or not; (C) The operator is controlling the robot in the virtual environment through an labyrinth; (D) The operator is interacting with the virtual environment to respond to warnings by touching one of the response cubes.

2.2 The Oddball Task

In the oddball-task, the subject had to respond to infrequent *task relevant* stimuli (*"targets"*, see Fig. 2 B) by touching one of three virtual cubes that were integrated into the virtual scenario as response targets for answering specific messages in a certain time frame (see Fig. 1 D). Besides the task relevant messages also frequent task irrelevant messages (*"standards"*, see Fig. 2 A) were presented to the operator but required no response. Due to the oddball design (Polich, 2007), i.e., the presentation of infrequent task relevant stimuli mixed with frequent task irrelevant stimuli, it was expected that P300 related brain processes (Kutas et al., 1977; Salisbury et al., 2001; Kirchner et al., 2009; Polich, 2007; Kirchner and Kim, 2012) will be evoked in case of recognized infrequent task relevant messages but not in case of frequent task irrelevant messages or task relevant but *not* recognized, i.e., missed, task relevant messages (*"missed targets"*). In recorded training data it was determined whether a target was recognized or not by the occurrence or missing of a response 10 sec after a target stimulus was presented.

2.3 The Operator Monitoring System

To support the operator in the scenario, an operator monitoring system (OMS) (Kirchner and Drechsler, 2013) was included into the setup. The purpose of the OMS was to monitor the operators cognitive state and the current state of scenario in order to adjust the course of events that were shown to the operator to minimize distraction of the operator and optimize her or his support by appropriate scheduling of messages.

The allowed response time was 2 sec in case that

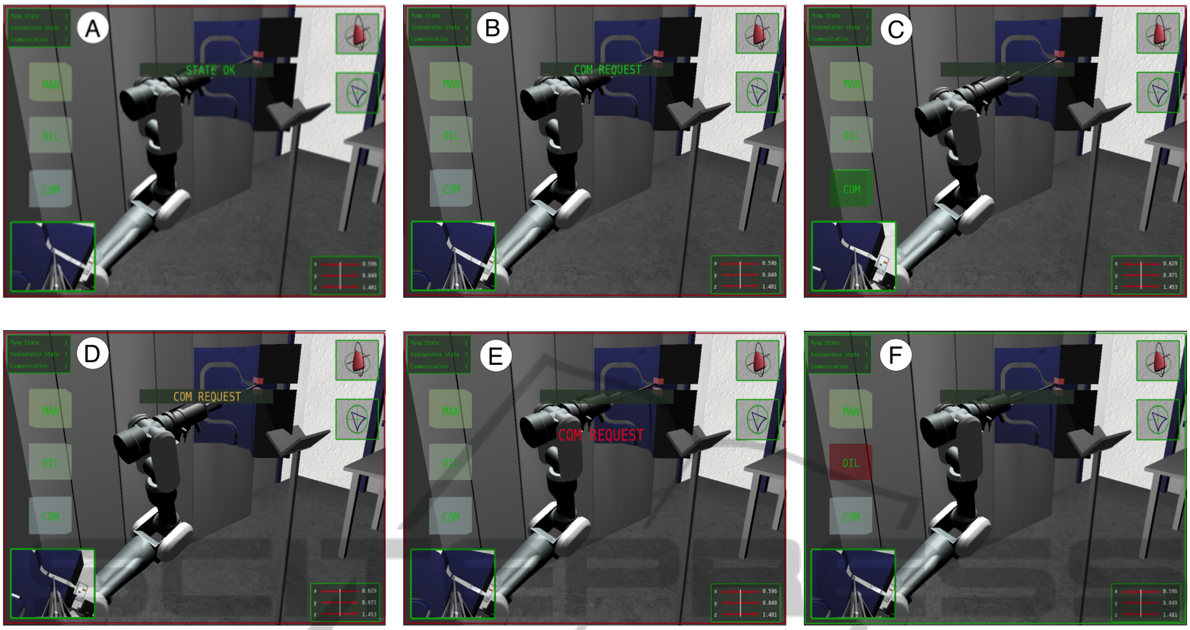

Figure 2: **The warnings and responses shown in the virtual immersion teleoperation operator monitoring scenario:** (A) The frequently shown *standard* marker has the text *STATE OK*. (B) One of the three possible *targets*. The possibilities are *MAN ENTERED*, *OIL TEMPERATURE* or *COM REQUEST*. (C) The operators response must be related to the warning. In this case, it is the cube with the label *COM*. In case of a correct response, it is highlighted in green. (D) If the operator did not respond in time, a repeated and highlighted second *target* is shown. (E) If the operator did not respond to the second target in time, a third obtrusive error message is shown. (F) In any case it might happen that the operator touches a cube with a label that does not correspond to the shown target. In this case, the cube is highlighted in red.

target recognition processes could *not* be detected after a target was presented (in case of *missed target*, see Fig. 1 B, first warning for "oil check" was missed) or to extend the allowed response time to 10 sec in case target recognition process were detected after a target was presented and recognized (in case of *target*, see Fig. 1 B, second warning for "oil check" was recognized).

By adapting the allowed response time in this manner allows to give the operator a longer time for responses in case he recognized the warning, which was especially relevant in case that the operator was in a rest position and was not able to respond. On the other hand a task relevant message could be repeated rather quick in case that target recognition processes could not be detected after a target message was presented.

To enable this adjustment of the virtual scenario by the OMS, we used machine learning techniques to detect the P300 ERP in the subjects EEG. Therefore, a classifier had to be trained to distinguish examples of the class *target* from examples of the class *missed targets* online. Since the operators were highly trained in the scenario they usually miss only a few target messages, hence the amount of missed target examples that could be recorded was very low. Thus, we used EEG activity evoked by irrelevant standard messages instead of EEG activity evoked by missed target messages during the training phase of the data processing to later distinguish between targets and missed targets.

3 METHODS AND EXPERIMENTAL PROCEDURES

The experiments were performed with three male subjects (age 27.33 ± 2.52), with a total of ten recording sessions.

3.1 Setup and Data Acquisition

The data was acquired with a 64-channel actiCap system and two amplifiers (both from Brainproducts, Munich, Germany) at 5 kHz sampling rate. Four electrodes (FC5, FC6, FT7, FT8) of the extended $10-20$ system were omitted to allow the HMD to be mounted. Thus, 60 channels were used for the recording.

The actual prediction system was active in four of ten sessions (online runs) and inactive in six out of ten sessions. Thus, the overall data consisted of 4 online and 6 pseudo-online sessions. Pseudo-online sessions consisted of three, and online sessions of four runs, with the forth run being the actual online run. Otherwise they were treated in the same way. A single run lasted in average about 13.64 ± 3.85 min.

The data acquired in the last run of each session was used for evaluation of our system, and the other sessions for training of the system. The pseudo-online sessions were analyzed offline after the experiments and are used here to provide a more comprehensive data basis.

The runs contained between 466 and 1553 (in average ≈ 865) *standards*, between 16 and 51 (in average ≈ 35) *targets*, and between 1 and 54 (in average ≈ 8) *missed targets*.

All processing was performed on equally-shaped windows of data with 1 s of duration, which were cutout and labeled according to the occurrence of a *standard*, *target* or *missed target*.

Between two runs there was a short break of $2-3$ min, except before the last run in the online sessions where the movement prediction system had to be trained and thus the break lasted around 10 min.

3.2 Processing Methods

We used our software pySPACE (Signal Processing And Classification Environment) (Krell et al., 2013a; Krell et al., 2013b) [1] for online and pseudo-online data analysis.

All windows were processed independently from each other. The data was preprocessed in several steps in order to extract the relevant features for the classifier. First, the data was standardized channel-wise by subtracting the mean signal value of the channel and divided by the standard deviation of the channel in the corresponding signal window. Next, a decimation with an anti-alias finite impulse response filter was performed to reduce the sampling rate of the data from 5 kHz to 25 Hz. This was followed by another band pass filter with pass band from 0.1 to 4.0 Hz.

Afterwards, the dimension of the data was reduced further in several steps. First, the xDAWN spatial filter (Rivet et al., 2009) was applied to reduce the 60 channels data to eight channels. The xDAWN spatial filter splits the data in noise and signal-plus-noise subspaces to, at the same time, extract meaningful signal information and reduce the dimension of the data. The number of eight channels was chosen based on previous experience with other experimental setups.

[1] available at http://pyspace.github.com/pyspace

We used straight lines, a special form of local polynomial features, as features for the classifier. The straight lines were fit channel-wise to segments of the data. Each segment lasted for about 400ḿs and adjacent segments overlapped by about 400ḿs. The slope of the lines were used as features. These were standardized again in the next processing step.

We used a soft-margin support vector machine (SVM) with a linear kernel for classification, where the complexity regularization hyperparameter was optimized using a grid search (tested values: $10^{-6}, 10^{-5}, \ldots, 10^{0}$) and an internal 5-fold cross validation. The data acquired in the training runs (see Sec. 3.1) were used for the training and hyperparameter optimization.

3.3 Evaluation

The data was evaluated with respect to the ability of a classifier to distinguish online and in single trial between EEG examples that contain patterns related to target recognition processes (that evoke a P300) and EEG examples that miss these patterns.

We expect that both, missed targets and standards would not evoke such patterns while targets do. Since enough training data was available we first analyzed how well a classifier trained on *standard* and *target* examples performs to distinguish between both classes (TS case).

Afterwards, we used a classifier that was trained on *standard* and *target* examples to distinguish between *target* and *missed target* examples (TM case) which was the relevant application case (see Sec. 2). To compare both results allows us to estimate how well classifier transfer performs in a demanding application teleoperation scenario.

4 RESULTS AND DISCUSSION

For the evaluation of the classification performance we use the Balanced Accuracy (BA), which is given by $\frac{1}{2}TPR + \frac{1}{2}TNR$, where TPR and TNR are the true positive and true negative rate, respectively. Accordingly, a BA score of 0.5 corresponds to random guessing, while a BA score of 1.0 would correspond to a *perfect* classifier. The BA is not affected by an unbalanced number of examples in each class, as it is the case here.

The obtained classification performance (as BA) for the different runs and sessions is shown in Table 1. The results correspond to the *single trial* detection of the P300 ERP.

Table 1: Classification performance for target vs. standard (TS case) and target vs. missed targets (TM case). The first 6 columns on the left contain the results for the *pseudo-online* sessions, the 4 columns on the right contain the results for the *online* sessions.

Evaluation	S1 R1	S1 R2	S2 R1	S2 R2	S3 R1	S3 R2	S1 $R_{ol}1$	S1 $R_{ol}2$	S1 $R_{ol}3$	S2 $R_{ol}1$
TS	0.84	0.97	0.96	0.95	0.88	0.86	0.91	0.88	0.85	0.93
TM	0.64	0.98	0.86	0.72	0.83	0.78	0.81	0.9	0.80	0.94

The average classification performance (as BA) in the TM case is 0.827 ± 0.103, which is slightly smaller as in the TS case (0.902 ± 0.048). It can be observed that in 8 of 10 sessions the difference of the BA is 0.1 or less, and in 5 of 10 sessions it is even 0.05 or less.

This shows that the achieved classification performance of our system works well for the P300 single trial detection for both the TM and TS cases despite the disturbance-prone setup. In addition, there is no remarkable difference between the TM and TS cases.

Furthermore, the difference between the BA score in the pseudo-online and online sessions is negligible (average classification performance TS case pseudo-online: $\approx 0.9089 \pm 0.057$, online: $\approx 0.893 \pm 0.037$; TM case pseudo-online: $\approx 0.803 \pm 0.110$, online $\approx 0.864 \pm 0.070$).

5 CONCLUSIONS AND FUTURE WORK

The presented results show that it is possible to detect the P300 in a complex and noisy application scenario where the operator of a robot has to perform a dual task, i.e., to teleoperate a robot and to respond to warnings. Furthermore, our results show that a classifier can be transferred between classes in case that both classes, here standard and missed targets, that miss a prominent pattern in the EEG signal, here the P300 ERP. This transfer works for most cases very well without a decrease of classification performance. However, in some cases the performance does decrease by a larger amount. Causes for this have to be investigated.

In future, we plan to improve the signal processing and pattern recognition methodology further to reduce the amount of required training data and to compensate for changes of the user itself, like fatigue, by applying online adaptation methods. In addition, we plan to use and evaluate our system in even more advanced and complex application scenarios, e.g., the supervision of operators that control several robots simultaneously.

ACKNOWLEDGEMENTS

This work was funded by the *Federal Ministry of Economics and Technology* (BMWi, grant no. 50 RA 1012 and 50 RA 1011).

REFERENCES

Farwell, L. A. and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr. Clin. Neurophysiol.*, 70(6):510–23.

Folgheraiter, M., Jordan, M., Straube, S., Seeland, A., Kim, S.-K., and Kirchner, E. A. (2012). Measuring the improvement of the interaction comfort of a wearable exoskeleton. *International Journal of Social Robotics*, 4(3):285–302.

George, L. and Lécuyer, A. (2010). An overview of research on "passive" brain-computer interfaces for implicit human-computer interaction. In *International Conference on Applied Bionics and Biomechanics ICABB 2010 - Workshop W1 "Brain-Computer Interfacing and Virtual Reality"*, Venice, Italy.

Guger, C., Harkam, W., Hertnaes, C., and Pfurtscheller, G. (1999). Prosthetic control by an EEG-based brain-computer interface (BCI). *5th European AAATE Conference*.

Haufe, S., Treder, M. S., Gugler, M. F., Sagebaum, M., Curio, G., and Blankertz, B. (2011). EEG potentials predict upcoming emergency brakings during simulated driving. *Journal of Neural Engineering*, 8(5):056001.

Iturrate, I., Montesano, L., and Minguez, J. (2013). Task-dependent signal variations in EEG error-related potentials for brain-computer interfaces. *Journal of Neural Engineering*, 10:026024.

Kim, S. K. and Kirchner, E. A. (2013). Classifier transferability in the detection of error related potentials from observation to interaction. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013, Manchester, UK, October 13-16, 2013*.

Kirchner, E. A. and Drechsler, R. (2013). A Formal Model for Embedded Brain Reading. volume 40, pages 530–540. Emerald Group Publishing Limited.

Kirchner, E. A. and Kim, S. K. (2012). EEG in Dual-Task Human-Machine Interaction: Target Recognition and Prospective Memory. In *Proceedings of the 18th Annual Meeting of the Organization for Human Brain Mapping*.

Kirchner, E. A., Kim, S. K., Straube, S., Seeland, A., Wöhrle, H., Krell, M. M., Tabie, M., and Fahle, M. (accepted 2013). On the applicability of brain reading for self-controlled, predictive human-machine interfaces in robotics. *PLoS ONE*.

Kirchner, E. A., Metzen, J. H., Duchrow, T., Kim, S. K., and Kirchner, F. (2009). Assisting Telemanipulation Operators via Real-Time Brain Reading. In Lohweg, V. and Niggemann, O., editors, *Proc. Mach. Learning in Real-time Applicat. Workshop 2009*, Lemgoer Schriftenreihe zur industriellen Informationstechnik, Paderborn, Germany.

Kirchner, E. A., Wöhrle, H., Bergatt, C., Kim, S.-K., Metzen, J. H., Feess, D., and Kirchner, F. (2010). Towards operator monitoring via brain reading – an EEG-based approach for space applications. In *Proc. 10th Int. Symp. Artificial Intelligence, Robotics and Automation in Space*, pages 448–455, Sapporo.

Krell, M. M., Straube, S., Seeland, A., Wöhrle, H., Teiwes, J., Metzen, J. H., Kirchner, E. A., and Kirchner, F. (2013a). pySPACE. https://github.com/pyspace.

Krell, M. M., Straube, S., Seeland, A., Wöhrle, H., Teiwes, J., Metzen, J. H., Kirchner, E. A., and Kirchner, F. (submitted 2013b). pySPACE a signal processing and classification environment in Python. *Frontiers in Neuroinformatics*.

Kutas, M., McCarthy, G., and Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science*, 197(4305):792–5.

Nijholt, A., Tan, D., Allison, B. Z., Del R Milan, J., and Graimann, B. (2008). *Brain-computer interfaces for HCI and games*, pages 3925–3928. ACM.

Pan, S. J. and Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol*, 118(10):2128–48.

Reuderink, B. (2008). Games and Brain-Computer Interfaces: The State of the Art. *WP2 BrainGain Deliverable HMI University of Twente September 2008*, pages 1–11.

Rivet, B., Souloumiac, A., Attina, V., and Gibert, G. (2009). xDAWN algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Transactions on Biomedical Engineering*, 56(8):2035–2043.

Salisbury, D., Rutherford, B., Shenton, M., and McCarley, R. (2001). Button-pressing Affects P300 Amplitude and Scalp Topography. *Clin Neurophysiol*, 112(9):1676–1684.

Seeland, A., Wöhrle, H., Straube, S., and Kirchner, E. A. (2013). Online movement prediction in a robotic application scenario. In *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clin. Neurophysiol.*, 113(6):767–91.

Zander, T. O., Kothe, C., Jatzev, S., and Gaertner, M. (2010). Enhancing human-computer interaction with input from active and passive brain-computer interfaces. *Brain-Computer Interfaces*, pages 181–199.