Striving for Better and Earlier Movement Prediction by Postprocessing of Classification Scores

Sirko Straube[1], Anett Seeland[2] and David Feess[2,3]
[1]*Robotics Group, Faculty of Mathematics and Computer Science, University of Bremen, Robert-Hooke-Str. 5, 28359 Bremen, Germany*
[2]*Robotics Innovation Center, German Research Center for Artificial Intelligence (DFKI GmbH), Robert-Hooke-Str. 5, 28359 Bremen, Germany*
[3]*Chair of Global Business, Faculty of Business Administration and Economics, University of Augsburg, Universitaetsstr. 16, 86159 Augsburg, Germany*

Keywords: EEG, LRP, Brain-computer Interface, Classification Score, Movement Prediction, Online Prediction.

Abstract: Brain-computer interfaces that enable movement prediction are useful for many application fields from telemanipulation to rehabilitation. Current systems still struggle with a level of unreliability that requires improvement. Here, we investigate several postprocessing methods that operate on the classification outcomes. In particular, the data was classified after preprocessing using a support vector machine (SVM). The output of the SVM, i.e. the raw score values, were postprocessed using previously obtained scores to account for trends in the classification result. The respective methods differ in the way the transformation is performed. The idea is to use trends, like the rise of the score values approaching an upcoming movement, to yield a better prediction in terms of detection accuracy and/or an earlier time point. We present results from different subjects where upcoming voluntary movements of the right arm were predicted using the lateralized readiness potential from the EEG. The results illustrate that better and earlier predictions are indeed possible with the suggested methods. However, the best postprocessing method was rather subject-specific. Depending on the requirements of the application at hand, postprocessing the classification scores as suggested here can be used to find the best compromise between prediction accuracy and time point.

1 INTRODUCTION

Movement prediction using the electroencephalogram (EEG) has a long standing history in the field of brain-computer interfaces (BCIs) since the discovery of readiness potentials (Kornhuber and Deecke, 1965; Libet et al., 1983), which build up long before the actual movement can occur. Since readiness potentials reflect preparatory activity and movement preparation can be aborted, these potentials can also disappear after a short build up without any movement occurring. However, the closer such a recorded potential gets to the actual movement, the stronger it is and the less likely will a prepared movement be cancelled (Fabiani et al., 2007, for a summary). When the movement is finally executed a corresponding motor potential can be recorded that reflects signalling to the muscles. For movement prediction, different signals have been applied, from the readiness potential itself over the lateralized readiness potential (LRP) which is closer to the movement and cannot easily be aborted (Blankertz et al., 2006), to specific frequency components in the EEG reflecting neural synchronization or desynchronization (Bai et al., 2011).

Movement prediction can be used as a powerful tool in various fields, with the most prominent being assistance during rehabilitation. Here BCIs predicting a movement can be used to close the gap between a patient's intention to move and the actual movement which can result in more intuitive responses of orthoses (Ahmadian et al., 2013; Kirchner et al., 2013; Kirchner and Tabie, 2013). Other fields include non-medical applications, e.g., during telemanipulation of a robotic device the user can be supported using a movement prediction based on EEG data (Folgheraiter et al., 2011; Folgheraiter et al., 2012). The idea is that the human operator experiences a smoother interaction with the telemanipulation device, which *knows* about an upcoming movement. As in the present study, the movement prediction is often based on the LRP.

Decisions in a movement predicting BCI come

Straube S., Seeland A. and Feess D..
Striving for Better and Earlier Movement Prediction by Postprocessing of Classification Scores.
DOI: 10.5220/0004632600130020
In *Proceedings of the International Congress on Neurotechnology, Electronics and Informatics* (NEUROTECHNIX-2013), pages 13-20
ISBN: 978-989-8565-80-8

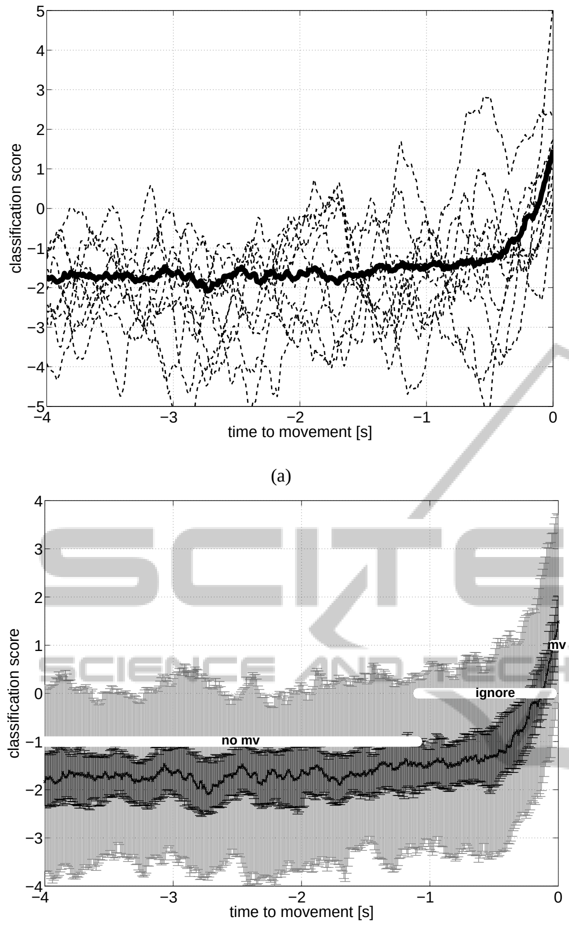


Figure 1: Example data from single subject prior to a movement. Depicted are 4 s of data with the movement onset at the very right (0 ms). **(a)** The bold black line shows the median of all 117 epochs. Dashed lines are 10 exemplary single trials. **(b)** Data of the same subject as (a) illustrated as 32/68 percentiles (black) and 5/95 percentiles (dark grey). The white line denotes time ranges where the data is labelled differently for evaluation: *no movement (no mv)* from -4000 ms to -1050 ms and *movement (mv)* from -50 ms to 0 ms. In between, data is ignored for true labels (see text).

from some kind of classifier which has to make the prediction. However, the output of the classifier is again noisy, so recent approaches try to apply a postprocessing to minimize classification errors (Lemm et al., 2004; Zhu et al., 2006; Solis-Escalante et al., 2008; Mohammadi et al., 2012). Here, we follow this rationale by applying simple online-capable functions to modify the classifier output according to knowledge about its progression. The scenario is the following: After processing, the classifier which is a support vector machine (SVM) assigns a value, the classification score, to each data instance (Vapnik, 1995). The range of these score values depends on the data at hand and on the classifier and can largely fluctuate as can be seen in Figure 1. A score of zero denotes the borderline between the two classes. The figure illustrates the high fluctuations in single trials and the consistent trend in the data: When the median score is considered, the score is constantly staying below zero, i.e., *no movement* is classified, until the scores rise approximately 500 ms before the movement and cross zero approximately at 250 ms before onset. The rise in classification scores before movement onset can consistently be observed across subjects. This means that the rise in score values alone may signal an upcoming movement so that the progression of the score values itself can be interpreted as being loosely correlated to the changes in movement probability.

The question now is whether we can use the knowledge about this rise in classification scores to make the prediction more stable and/or predict the upcoming movement earlier. In trying to answer this question we were seeking for a postprocessing method that dampens fast fluctuations in classification scores and stabilizes long rises. To this aim, we applied several methods that modify the current classification scores by taking into account previous scores with a certain weight (see Section 2).

To summarize, if an LRP can be detected by high levels of the classification score, it could potentially just as well be predicted earlier by detecting the rise that leads to that elevated level. In the following we will describe the postprocessing methods that we have applied. After a description of the experimental data used, the results will be presented and discussed.

2 POSTPROCESSING METHODS

From the perspective of a movement prediction application it is most desirable to perform robust, binary decisions: A movement will either occur or it will not. This decision should be made as reliably and early as possible. From the large margin classification perspective, this means that the classification score S_t at some point in time t would have to be compared against some threshold b so that a movement *mv* is predicted when

$$\text{mv iff} \quad S_t \geq b. \tag{1}$$

Yet, as illustrated in Figure 1, the score sometimes suddenly crosses the threshold when the actual movement is still far away, but then only for a short time. This behaviour hinders reliable prediction when it is purely based on the raw value of S_t crossing b. Looking at the average score progression over time reveals a continuous rise of the score values before the actual movement. Here, we exploit this systematic behaviour to find a function F that is able to generate better movement predictions based on past values of S, such that

$$\text{mv iff} \quad F(S_t, S_{t-1}, \dots, S_{t-(k-1)}) \geq b_F \tag{2}$$

for some specific threshold b_F. k is defined as the number of scores that are used in F with the current score being at $k = 1$. In principle, there are no constraints on the functional form of F.

In the present study we apply weights to the current and previous $k-1$ classification outcomes to transform the current score S_t. These weights decay with the number of steps looked into the past. We also followed an alternative approach by transforming the current score with the average slope of the past samples. A detailed description is given in Subsection 2.2. Both types of functions (weighting and slope approach) can be expressed as

$$F(S_t, S_{t-1}, \dots, S_{t-(k-1)}) = \\ w_1 S_t + w_2 S_{t-1} + \dots + w_k S_{t-(k-1)} \tag{3}$$

with some predefined weights w. With this methodology we try to boost the score value when previous scores were similar in value and at the same time penalize scores when previous ones showed a completely different trend. The approaches are described in more detail in the following.

2.1 Fixed Weighting

In this set of functions the weights are generated by very simple functions, each of which assigns a high weight to the most current classification score sample, and decreasing weights to older samples.

The functions used are depicted in Fig. 2. All functions have in common that the weights add up to one. The coefficients for the `uniform`, `linear`, `square`, and `cubic` method are all generated by evaluating

$$w_\tau = \frac{\tau^p}{\sum_{i=1}^{k} i^p}, \quad \tau \in \{1, \dots, k\}, \tag{4}$$

respectively, with k the number of coefficients used and the exponent p according to the corresponding function type. The `exp` coefficients are accordingly calculated as

$$w_\tau = \frac{\exp \tau}{\sum_{i=1}^{k} \exp i}, \quad \tau \in \{1, \dots, k\}. \tag{5}$$

Besides these rather universal functions for choosing the weight we added two variants where we explicitly forced the current value to have a much higher

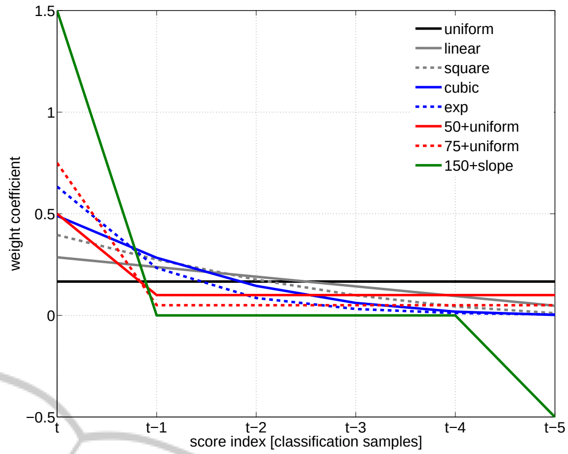


Figure 2: Comparison of the functions used for classification score postprocessing using $k = 6$ coefficients, i.e., the current score and 5 instances back in time.

weight than the scores corresponding to the previous instances, since the idea behind the postprocessing was exactly this: Transform the current score with its history to weaken fast fluctuations and strengthen longer trends. Again, the weights were set so that they add up to one. In the `X+uniform` method, the first coefficient gets assigned a weight of $X\%$. The remaining weight of $[1-(X/100)]$ is then equally distributed across the remaining coefficients.

2.2 Slope Approaches

Since the objective is to identify a rise in the classification score progression over time we also looked at modifications of the score value using local slopes or averaged slope over the last k samples (i.e., the current sample and $k-1$ instances back in time). Considering two samples, a local slope ΔS_t^1 can be computed as

$$\Delta S_t^1 = \frac{S_t - S_{t-1}}{t-(t-1)} = S_t - S_{t-1}. \tag{6}$$

Therefore, the average slope ΔS_t^{k-1} over k samples is

$$\Delta S_t^{k-1} = \frac{1}{(k-1)} \sum_{i=1}^{k-1} (S_{t-i+1} - S_{t-i}). \tag{7}$$

which is a telescope sum and boils down to

$$\Delta S_t^{k-1} \propto (S_t - S_{t-(k-1)}). \tag{8}$$

The corresponding weighting coefficients for this postprocessing are then

$$w_1 = 1, \; w_k = -1, \; w_\tau = 0 \; \forall \tau \notin \{1, k\}, \\ \text{or} \quad w = (1, 0, \dots, 0, -1). \tag{9}$$

In pilot experiments (not shown) this `slope` method was tested and performance levels were consistently far below the performance obtained without any postprocessing. Due to these performances losses of at least 0.15 points of balanced accuracy (BA, see Section 4.1) and in worst cases a performance around the probability of guessing this method was skipped for the current study.

Nevertheless, since we were looking for stabilizing a slope, we chose another promising and simple variant. Instead of using only the slopes, we modulate the current score with the slope approach in a 2:1 fashion (score:slope), so that we obtain a weight vector w of

$$w = (1.5, 0, \ldots, -0.5). \tag{10}$$

In other words, in this approach we take the current score value with 100% and add the slope weighted with 0.5. This variant is called `150+slope`.

3 DATA & PREPROCESSING

The data used for evaluation has been described in detail previously (Kirchner and Tabie, 2013; Tabie and Kirchner, 2013). Originally, muscle activity has been recorded simultaneously with the EEG. Here, evaluation has been restricted to EEG data.

3.1 Experimental Data

Eight right-handed male subjects (age: 29.9 ± 3.3 years) participated in the study. They gave written consent to participate and could abort the experiment at any time. The study was conducted in accordance with the Declaration of Helsinki. The subjects were sitting in a comfortable chair in front of a table with a monitor showing a fixation cross and giving occasional feedback. They executed self-paced, intentional movements with their right arm by releasing a button and pressing another one situated 30 cm to the right. A resting period of 5 s between movements had to be performed for a movement to be counted as valid. Subjects were not informed about this time constraint, instead negative feedback was provided (a red circle around the fixation cross) when they performed a movement too quickly after another. In each session 120 correctly performed movements were recorded, divided into 3 runs (40 movements per run).

3.2 Preprocessing

The EEG was acquired with 5 kHz, filtered between 0.1 Hz to 1 kHz using the BrainAmp DC amplifier [Brain Products GmbH, Munich, Germany]. Recordings were performed using a 128-channel (extended 10-20) actiCap system (reference at FCz). Electrodes I1, OI1h, OI2h and I2 were used for electrooculography and thus not placed on the scalp. For detection of the physical movement onset a motion capturing system consisting of 3 cameras (ProReflex 1000; Qualisys AB, Gothenburg, Sweden) was used at 500 Hz. After synchronization of the two data streams, the movement onsets were marked in the EEG.

Preprocessing was performed on overlapping windows of 1 s length cut every 10 ms in a range from -4000 ms to 0 ms before a movement. Consequently, a total of 401 score values were computed per executed movement. Data were standardized channel-wise (subtraction of mean and division by standard deviation) and decimated to 20 Hz. Next, a FFT band-pass filter with a pass band of 0.1 to 4 Hz was applied. Since the prediction should be based on the most recent data, we proceeded with the last 200 ms of each window that were processed by an xDAWN spatial filter (Rivet et al., 2009) with 4 channels retained. For feature extraction, raw voltage values were used, standardized (mean zero, variance one) and classified by a SVM (Chang and Lin, 2011) with linear kernel.

For trainable components in the signal processing chain (xDAWN, feature normalization and SVM) windows ending at -100 and 0 ms were labeled as *movement*. Training windows for *no movement* originated from non-overlapping windows (1 s length) that were continuously cut from the data stream, if no movement occurred 1 s before and 2 s after this window. In addition, a parameter optimization for the complexity parameter of the SVM was performed using a grid search (tested values: $10^0, 10^{-1}, \ldots, 10^{-6}$). A 3-fold cross-validation, one fold corresponding to one experimental run, was applied and classifier scores were stored for both, training and test data.

4 EVALUATION

As the aim is *to detect movements more accurately and/or earlier*, there are basically two criteria for a good postprocessing. One is the detection accuracy, the other the time point of detection. Both are considered for evaluation.

4.1 Movement Detection Accuracy

The prediction of unique events comes along with unbalanced proportions of the two classes *no movement* and *movement*, i.e., class instances of data containing the LRP (in our case) will be underrepresented. The evaluation of the movement detection accuracy has to take this into account. Thus, the simple accuracy is misleading (Kubat et al., 1998, for discussion), so a metric is required which is insensitive to imbalanced classes.

One of the most intuitive measures existing in such a case is the *balanced accuracy* (BA) which is defined as the mean of true positive rate (TPR) and true negative rate (TNR):

$$BA = \frac{TPR + TNR}{2}. \tag{11}$$

One of the challenges here is to define a ground truth of when the relevant signal (i.e., the LRP) is actually present in the data. While we can postulate that there must be an LRP prior to each movement, we still do not know the precise onset of this signal. To cope with this issue and thereby get unambiguously labelled data for evaluation, we split the time before a movement into three phases (compare Figure 1), a *no movement* phase from -4000 ms to -1050 ms, a *movement* phase from -50 ms to 0 ms, and the phase in between (-1050 ms to -50 ms) where the data is ignored and not labelled at all. With this approach we obtain a clear labeling in phases where we are sure that the relevant signal is indeed contained in the data. This signal is, of course, also present in the ignored time range, but since we do not know the exact onset this range is skipped. In the actual application where no data are skipped, a movement is predicted whenever the classifier score crosses the threshold.

4.2 Time Point of Detection

The onset of the signal related to movement (here, the LRP) occurs at an unknown point in time before the actual movement. This transition out of noise is typical for event-related potentials and it is reflected in the rise of classification scores that we intend to stabilize with the postprocessing approaches introduced here. Concerning the application, i.e., the prediction of an upcoming movement, the exact time point is of less importance than a *reliable and stable* prediction by the classifier. For this, it remains to define when exactly we consider the LRP as *detected*—the classification score might at any time rise over a given threshold for a short period of time due to noise. For the same reason, the score might fall below the threshold for some samples although the LRP has—supposedly—already been correctly detected. To make sure that we base our evaluation on a stable prediction, the *LRP onset* was defined as the point in time where the classification scores do not drop below the threshold for N predictions. This point was found by going back in time from the actual movement onset until the first time where this criterion was not met. With this method, the LRP onset used for evaluation is then defined as the first score sample crossing the threshold after the set of samples staying below threshold for N predictions.

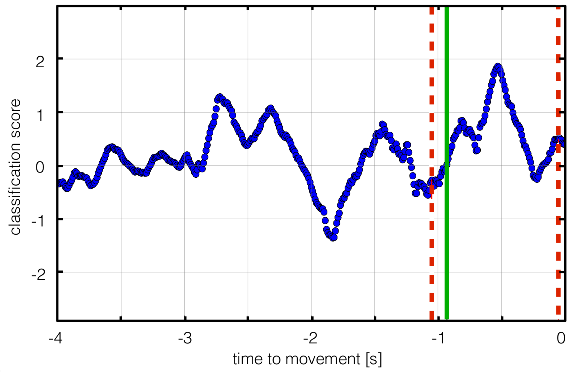
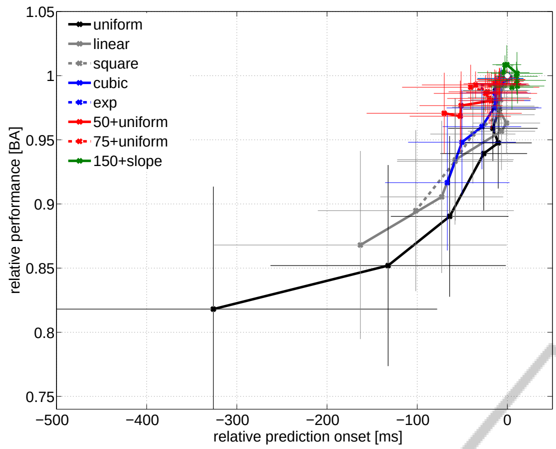


Figure 3: Single trial example for determination of LRP onset with $N = 10$ specifying the number of samples a false negative classification is tolerated. The dashed red lines delimit the feasible (transition) area, the solid green line indicates the detected LRP onset. Setting $N = 10$ provokes that the small dip around -200 ms is ignored.

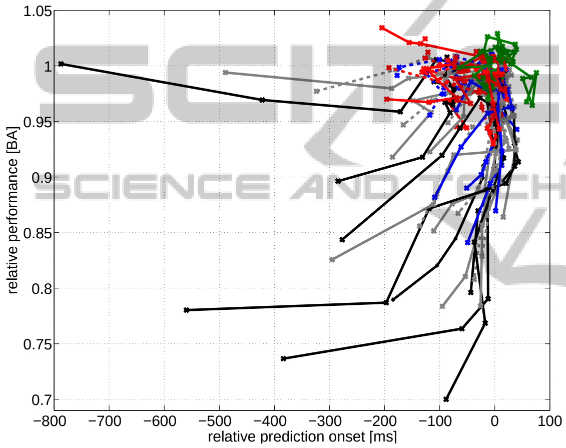
The choice of N depends on the level of noise, on the classification scores, on the sampling rate, and on the characteristics of the signal applied for movement prediction. Here, the relevant signal has a length of approximately 1 s (Fabiani et al., 2007), so we chose $N = 10$ as a good compromise between robustness (higher N) and reliability (lower N), i.e., we tolerate false classifications during periods shorter than 100 ms. Increasing the robustness here means to allow an earlier estimation of the time point of detection, because fluctuations in the score progression are more and more ignored with increasing N. On the contrary, a decrease of N increases the reliability, because fewer classifications of *no movement* can occur after the estimated time point of detection, but this comes at the cost of a higher sensitivity to outliers. To give an alternative view on the value of N: Setting $N = 10$ in our data means that the movement onset is defined as the first score sample crossing the threshold *after* a 100 ms window without any predicted movement (viewed backwards from the actual movement onset). The approach is illustrated in Figure 3.

4.3 Evaluation Procedure

For each subject and cross-validation fold two data sets exist: one training data set (80 movements) and one test data set (40 movements). The training set is the one used to train the classifier producing the classification scores. Due to the fact that the postprocessing methods introduced here change the absolute value of the classification score, the score thresholds (transition from one class to the other) were re-adjusted for each method, respectively, using the



(a)



(b)

Figure 4: Performance changes (BA and onset time) of postprocessing methods. The results are illustrated relative to the case where the scores were not further processed, illustrated as a white dot at (0,1). Each data point corresponds to a different $k \in \{1,2,4,8,12,16,20,40,60,100\}$ (see text). In (a) the grand average results over all subjects are shown as mean and standard deviation for each k, respectively. Thus, there is one line for each method applied. In (b) the same results are shown for all subjects, separately.

training data, before evaluation was performed on the test data. The results presented in the following show the performance in terms of detection accuracy and time point of detection on the test data.

5 RESULTS & DISCUSSION

All investigated methods introduced in Section 2 and illustrated in Figure 2 were tested for different values of the parameter $k \in \{1,2,4,8,12,16,20,40,60,100\}$ which is the number of scores used with the respective method. Since a key motivation for the current work was to modify the current score by the values of the neighbouring scores, we chose a finer granularity for sampling near the current score. Without postprocessing, the movements were predicted on average over all subjects 180 ms before the movement onset with a balanced accuracy of 0.8. Since we were interested in performance improvements concerning these two measures, the results are illustrated as relative changes according to these *reference performances* in Figure 4 (a). The figure shows the average result for all methods applied for the two criteria detection accuracy and time point of detection (see Section 4). Here each data point corresponds to a particular value for k, with $k = 100$ being the first data point on the lower left of the plot and the next smaller value for k being the next on the connecting line. Finally, all methods meet at $k = 1$ on the upper right at a relative prediction onset of 0 ms and a relative performance of 1, because all of these have an identical weight vector of $w = (1)$. This point (highlighted as white spot in Figure 4) with $k = 1$ is equal to the reference performance without any postprocessing.

The results in Figure 4 (a) indicate that the performance obtained when using the raw score values was already on a high level regarding the time point of detection *and* the classification accuracy. In the figure, a postprocessing method outperforming this reference would be on the upper left relative to this point. This we observed only for the slope approach `150+slope`, where we found configurations for $k \in \{2,4\}$ that revealed a slight improvement on average in both, accuracy and time point.

From the figure, it is far more apparent that most methods enabled an earlier prediction of the movement on the cost of (mostly) slight performance drops. In the most extreme case for the `uniform` approach with $k = 100$, this means more than 300 ms earlier prediction at a loss of 18% of the initial performance. Overall improvements in classification accuracy on the average level were only revealed for the `150+slope` approach.

The reason for the large standard deviations depicted in Figure 4 (a) is disclosed by illustration of the single-subject results in Figure 4 (b). The benefit of the applied method strongly differed between subjects, so that the results in Figure 4 (a) only show the rough trend. On the single-subject level we observed slight improvements for both, time point and accuracy. However, the *best* method was subject-specific. Again, most extreme differences where achieved using the `uniform` approach: Using $k = 100$ we could detect the movement nearly 800 ms earlier without any performance loss for one subject, while the same configuration resulted in only 100 ms earlier detection at a loss of 30% of the initial performance for another. In the analysis, especially subjects with a worse performance on the raw scores could benefit from the postprocessing.

To summarize, the postprocessing methods presented here provide a tool to modify mainly the earliness of the prediction and to a little extent the classification accuracy. The `150+slope` method with $k \in \{2,4\}$ worked best on the average level and enhanced both, time point of detection and accuracy. On the single-subject level, the individual best method differed, so that the spectrum presented here can serve as a general framework to *adjust* the movement prediction according to the respective application and/or the data of the individual subject.

6 CONCLUSIONS

Without any postprocessing, the classification of each window is performed independently of the neighbouring windows. However, we can see in the distributions of these classification outcomes that they intrinsically carry information about the probability of an upcoming event, like the rise in scores illustrated in Figure 1. Here, we use simple methods that can easily be applied during online movement prediction to make use of this knowledge and stabilize a single classification by the surrounding ones. The methodology introduced here can be used as a tool to improve classification outcomes.

Since the performance *and* the time point of prediction can be both equally relevant for an application, we considered these two measures together. Taking these, we could show that the applied methods succeed for individual subjects in improving the accuracy and/or time point of prediction, although we could not find one straightforward solution in the current study for all subjects investigated. For most methods we found a trade-off between these two metrics. This means for the application of such a movement predicting system, that one can indeed enhance the system, but has to carefully chose the postprocessing method according to the requirements of the intended application.

On average we could observe an improvement of both, time point of detection and accuracy, using the `150+slope` method with small values of k. However, we found the most pronounced effects on the single-subject level: the proposed methods performed individually different. In an application, such a high subject-specificity can be dealt with using two approaches: either extra calibration time is used to find the best individual method, or the prediction itself is integrated in the application in a way that is flexible or robust enough to make use of the possible benefits illustrated here. Since the predictions obtained without postprocessing can serve as a reliable fallback option, this could be realized, e.g., by using a number of the proposed approaches on top and making the final prediction from the ensemble. More in general, although the finding of a subject specificity is consistent with results from other postprocessing methods (Mohammadi et al., 2012), it should be helpful to reveal the particular origins of these effects in the score progression to develop a method that generalizes better. So far, existing postprocessing methods operate rather blindly on the data which may cause the individual differences.

While most research is dedicated to improvement of the classifier and/or preprocessing algorithms, the idea of postprocessing of classification outcomes as such is not completely new. Techniques for incorporating preceding probabilities to enhance the current prediction have been proposed (Lemm et al., 2004; Zhu et al., 2006), but not been evaluated in the way we did in the present study. Therefore and since SVM scores do not directly represent probabilities like in a Bayesian framework, a direct comparison with the methods we proposed is difficult. However, from the technical point of view all of these methods have in common that they actually manipulate single prediction outcomes making use of the individual prediction history. Other techniques exist for postprocessing that rather operate on the global level by changing the decision criterion of the classifier or using additional thresholds. Here, threshold selection, dwell time optimization or debiasing of the score time course have been proposed (Solis-Escalante et al., 2008; Mohammadi et al., 2012). Due to their different nature, these techniques can be easily combined with what we proposed here, as we already implicitly did by including threshold optimization (see Section 4.3) and selecting a stability criterion of 100 ms ($N = 10$; see Section 4.2 and Figure 3), which can be interpreted as a dwell time.

With the approach outlined here, other and more complex algorithms can of course be used, although they might have the possible drawback of being too computationally complex for an online predicting system. Generally, the methods applied here are not specific for the context of movement prediction, so they can be used in any context where such postprocessing may be helpful.

ACKNOWLEDGEMENTS

This work was supported by the German Bundesministerium fur Wirtschaft und Technologie (BMWi, grant FKZ 50 RA 1012 and grant FKZ 50 RA 1011). The authors like to thank Marc Tabie for providing us with the evaluation data.

REFERENCES

Ahmadian, P., Cagnoni, S., and Ascari, L. (2013). How capable is non-invasive EEG data of predicting the next movement? A mini review. *Frontiers in Human Neuroscience*, 7:124.

Bai, O., Rathi, V., Lin, P., Huang, D., Battapady, H., Fei, D.-Y., Schneider, L., Houdayer, E., Chen, X., and Hallett, M. (2011). Prediction of human voluntary movement before it occurs. *Clinical Neurophysiology*, 122(2):364–372.

Blankertz, B., Dornhege, G., Lemm, S., Krauledat, M., Curio, G., and Müller, K. (2006). The Berlin Brain-Computer Interface: machine learning based detection of user specific brain states. *Journal of Universal Computer Science*, 12(6):581–607.

Chang, C.-C. and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

Fabiani, M., Gratton, G., and Federmeier, K. D. (2007). Event-related brain potentials: Methods, theory, and applications. In Cacioppo, J., Tassinary, L. G., and Berntson, G. G., editors, *Handbook of Psychophysiology*, pages 85–119. Cambridege University Press, Cambridge [u.a], 3rd edition.

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.

Folgheraiter, M., Kirchner, E. A., Seeland, A., Kim, S. K., Jordan, M., Wöhrle, H., Bongardt, B., Schmidt, S., Albiez, J., and Kirchner, F. (2011). A multimodal brain-arm interface for operation of complex robotic systems and upper limb motor recovery. In Vieira, P., Fred, A., Filipe, J., and Gamboa, H., editors, *In Proceedings of the 4th International Conference on Biomedical Electronics and Devices (BIODEVICES-11)*, pages 150–162, Rome. SciTePress.

Kirchner, E. A., Albiez, J., Seeland, A., Jordan, M., and Kirchner, F. (2013). Towards assistive robotics for home rehabilitation. In Chimeno, M. F., Solé-Casals, J., Fred, A., and Gamboa, H., editors, *In Proceedings of the 6th International Conference on Biomedical Electronics and Devices (BIODEVICES-13)*, pages 168–177, Barcelona. SciTePress.

Kirchner, E. A. and Tabie, M. (2013). Closing the gap: combined EEG and EMG analysis for early movement prediction in exoskeleton based rehabilitation. In *Proceedings of the 4th European Conference on Technically Assisted Rehabilitation - TAR 2013*.

Kornhuber, H. H. and Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, 284(1):1–17.

Kubat, M., Holte, R. C., and Matwin, S. (1998). Machine learning for the detection of oil spills in satellite radar images. *Machine Learning*, 30(2-3):195–215.

Lemm, S., Schäfer, C., and Curio, G. (2004). BCI competition 2003–data set III: probabilistic modeling of sensorimotor mu rhythms for classification of imaginary hand movements. *IEEE Transactions on Biomedical Engineering*, 51(6):1077–80.

Libet, B., Gleason, C. A., Wright, E. W., and Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential) the unconscious initiation of a freely voluntary act. *Brain*, 106(3):623–642.

Mohammadi, R., Mahloojifar, A., and Coyle, D. (2012). A combination of pre- and postprocessing techniques to enhance self-paced BCIs. *Advances in Human-Computer Interaction*, 2012:3:1–3:10.

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.

Solis-Escalante, T., Müller-Putz, G., and Pfurtscheller, G. (2008). Overt foot movement detection in one single laplacian EEG derivation. *Journal of Neuroscience Methods*, 175(1):148–153.

Tabie, M. and Kirchner, E. A. (2013). EMG onset detection – comparison of different methods for a movement prediction task based on EMG. In Alvarez, S., Solé-Casals, J., Fred, A., and Gamboa, H., editors, *In Proceedings of the 6th International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS-13)*, pages 242–247, Barcelona. SciTePress.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer New York, Inc., New York, NY, USA.

Zhu, X., Wu, J., Cheng, Y., and Wang, Y. (2006). GMM-based classification method for continuous prediction in brain-computer interface. In *Proceedings of the 18th International Conference on Pattern Recognition - Volume 01*, ICPR '06, pages 1171–1174, Washington, DC, USA. IEEE Computer Society.