Published in final edited form as:

J Neurosci Methods. 2011 July 15; 199(1): 103–107. doi:10.1016/j.jneumeth.2011.04.037.

SHOULD THE PARAMETERS OF A BCI TRANSLATION ALGORITHM BE CONTINUALLY ADAPTED?

Dennis J. McFarland, William A. Sarnacki, and Jonathan R. Wolpaw

Laboratory of Neural Injury and Repair, Wadsworth Center, New York State Department of Health, Albany, New York 12201-0509

Abstract

People with or without motor disabilities can learn to control sensorimotor rhythms (SMR) recorded from the scalp to move a computer cursor in one or more dimensions or can use the P300 event-related potential as a control signal to make discrete selections. Data collected from individuals using an SMR-based or P300-based BCI were evaluated offline to estimate the impact on performance of continually adapting the parameters of the translation algorithm during BCI operation. The performance of the SMR-based BCI was enhanced by adaptive updating of the feature weights or adaptive normalization of the features. In contrast, P300 performance did not benefit from either of these procedures.

INTRODUCTION

Many people with severe motor disabilities require alternative methods for communication and control. Numerous studies over the past two decades indicate that scalp-recorded EEG activity can be the basis for nonmuscular communication and control systems, commonly called brain–computer interfaces (BCIs) (Wolpaw et al., 2002). EEG-based communication systems extract specific features of EEG activity and translate them into control signals. Some systems use as features EEG components in the frequency domain that are spontaneous in the sense that they are not dependent on specific sensory events (e.g., Wolpaw and McFarland (1994, 2004)). Others use as features EEG potentials that are evoked by stereotyped stimuli (e.g., Farwell and Donchin 1988; Sellers et al., 2006).

Effective BCI operations depend on use of appropriate methods for recording brain signals, extracting features from these signals, and translating these features into device commands (Wolpaw et al 2002). Feature selection and translation is extremely important, and has been the subject of numerous studies and data competitions (e.g. Blankertz et al (2006), see McFarland et al (2006a) for a review). An effective method selects and weights relevant features in proportion to the information they contain about the user's desired outcome. Many kinds of methods, linear as well as nonlinear, are possible (e.g. Muller et al (2003)).

There are at least three distinct approaches to improving BCI operation (McFarland et. al., 2006b). The first is to use machine learning to improve feature selection and translation. The

© 2011 Elsevier B.V. All rights reserved.

Corresponding author: Dennis J. McFarland, Wadsworth Center, PO Box 509, Empire State Plaza, Albany, NY 12201-0509, (518) 473-4680, mcfarlan@wadsworth.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

second is to train the user to provide stable and effective features. The third approach is a combination of the first two; it views BCI operation as the joint product of two adaptive controllers, the user and the BCI system, and seeks to optimize their interaction. The first and third approaches emphasize the importance of adapting the feature selection and translation process.

The value of adapting feature selection and translation may depend on the kind of features that the BCI uses (Sellers et al., 2007). For example, user training can play an important part in operating a sensorimotor rhythm (SMR)-based BCI (e.g., Wolpaw and McFarland, 2004). Thus, it may be useful for feature selection and translation to adapt as learning occurs. In contrast, the P300-based BCI may operate effectively for lengthy periods after an initial adaptation to user-specific characteristics (Krusienski et al., 2008).

Each step in the process of feature extraction and translation may involve estimating parameters that are derived from the observed data. When these data are not stationary, some form of adaptation may optimize BCI performance (McFarland et al., 2006a). Furthermore the criteria for adaptively estimating these parameters may vary with the parameter in question (McFarland et al., 2006b). For example, tracking variations in the mean of the signal features can be computed from the data in an unsupervised manner (e.g. Ramoser et al., 1997). In contrast, adaptation of the weights for a model that translates EEG features into cursor movement represents supervised learning based on knowledge of intended trial outcomes (e.g., Fabiani et al., 2004).

The value of ongoing adaptation of feature selection and translation has been noted for some time now (e.g., Wolpaw et al., 1991), and several adaptive systems have been described (e.g., Krausz et al. (2003), Sykacek et al (2004), Ince et al (2006), Vidaurre et al (2006), Shenoy et al. (2006)). However, firm empirical support for ongoing adaptation, including evidence that is important for long-term BCI operation, is generally lacking. In addition, it is not clear which BCI systems are most likely to benefit, nor which parameters should be adapted or how this should be accomplished.

The present study set out to assess the ability of ongoing adaptation to improve the long-term performance of an SMR-based BCI and a P300-based BCI. In separate offline analyses of data gathered over multiple daily sessions, it estimated the impact on performance of adaption of feature weights and adaptive normalization of features. While the results are informative, they may not generalize completely to online operation.

METHODS

SMR-Based BCI Operation and Data Acquisition

The SMR users were 5 individuals (3 women and 2 men, 27–55 years old) who had no prior BCI experience. They learned a two-target cursor movement task according to our standard protocol (McFarland et al, 2005).

The BCI user sat in a reclining chair facing a video screen and remained motionless. BCI operation and data collection were supported by the general-purpose BCI software platform BCI2000 (Schalk *et al* 2004) in conjunction with a 64-channel SA Instrumentation amplifier and a Data Translation DT-3003 64 channel A/D board. EEG was recorded from 64 scalp locations (Sharbrough *et al* 1991) by 9-mm tin electrodes embedded in a cap (Electrocap International) and referenced to an electrode on the right ear, and was digitized at 160 Hz and stored for later analysis.

Each trial began when a target consisting of a vertical bar half the height of the screen appeared along either the top half or the bottom half of the right edge of the screen. One second later a cursor appeared in the middle of the left edge of the screen and began moving horizontally to the right at a constant rate and moving vertically under control by the user's EEG. The cursor took four seconds to move across the screen, and the user's task was to move the cursor vertically so that it hit the target when it reached the right edge (chance was 50%). During the 1-sec feedback interval, the target turned yellow if it had been hit by the cursor; otherwise the screen went blank. The next trial began after a 1.5-sec. intertrial interval. Each session consisted of eight 3-minute runs separated by 1-min rests. The user's first six sessions provided the data (i.e., about 160 trials/session or 1080 total trials) for the current offline analyses.

The features analyzed were the magnitude of 3-Hz wide spectral bins computed from a 16th order autoregressive model applied to 400 msec data segments (McFarland and Wolpaw, 2008). Bins centered at 9, 12, 21 and 24 Hz were computed from Laplacian derivations (McFarland et al 1997) at electrode locations C3, C4, CP3 and CP4 (after Sharbrough et al., 1991). This provided 16 features for the translation algorithm.

P300-Based BCI Operation and Data Acquisition

The P300 users were 5 individuals (3 men and 2 women, 23–53 years old) who had no prior BCI experience. They performed P300-based BCI copy-spelling with a 6x6 matrix according to our standard protocol (Krusienski et al., 2008). The first session consisted of copy-spelling without feedback. This session provided the data for developing the classifier that was used to provide feedback (i.e., which letter was selected) during the next five sessions.

BCI operation and data collection were supported by the general-purpose BCI software platform BCI2000 (Schalk et al 2004) in conjunction with a g-Tec g. USBamp 16 channel data acquisition system. EEG was recorded from 16 scalp locations (after Krusienski et al., 2008) by 9-mm tin electrodes embedded in a cap (Electrocap International) and referenced to an electrode on the right ear, and was digitized at 256 Hz and stored for later analysis.

Briefly, the user sat upright in front of a video monitor and viewed the 6x6 matrix of letters and other items. All data were collected in the copy-speller mode: the letters for selection were presented on the top left of the video monitor (above the matrix) and the letter currently specified for selection was in parentheses at the end of the letter string. The task was to attend to the specified letter in the matrix and silently count the number of times that letter flashed, until a new letter was specified for selection. The rows and columns of the matrix flashed for 125 ms with 62.5 ms between flashes. One complete cycle of six row and six column flashes constituted a sequence and 15 sequences constituted a single trial (i.e., a single letter selection). Specifically, classification (i.e., letter selection) was performed after every row and column has been intensified 15 times. Each session consisted of nine 3-min runs; each run comprised the spelling of a word or series of letters chosen by the investigator. This set of letters spanned the letters contained in the matrix and remained stable across each person's six sessions. The features for offline analyses were the 20 successive 31.25-msec averages in the first post-flash 625-msec EEG segments from electrodes Fz, Cz, Pz, PO7 and PO8. This resulted in 100 features for classification. Each session provided 6300 non-target and 1260 target EEG segments. Since waveforms associated with multiple stimuli were not averaged the performance is somewhat lower than typically reported in the literature. However this procedure avoids ceiling effects.

Adaptation of Feature Weights

In both the SMR-based BCI and the P300-based BCI, the output (i.e., vertical cursor movement for the former and letter selection for the latter) was determined by a linear classifier in which the independent variables were the chosen EEG features and each feature was weighted as specified by a least-squares discriminant analysis applied to a particular set of data.

In offline analysis of each subject's data for sessions 2–6, we compared the BCI performances when four different sets of training data were used to determine the weights. These four sets were: *Cross*: weights were obtained by leave-one-out crossvalidation with data from the current session; *First*: weights were based on data from session 1; *Last*: weights were obtained with data from the immediately preceding session; and *All*: weights were obtained with data from all of the preceding sessions. These offline analyses were done with the SAS Discrim procedure.

The Cross method is a standard offline analysis option that could not be used online. The First method is a standard nonadaptive method that could be used online. The Last and All methods are adaptive methods that could be used online.

Adaptive Normalization

We also assessed the impact on SMR-based and P300-based BCI performances of normalizing the session 2–6 data prior to leave-one-out cross-validation of the pooled data with the SAS Discrim procedure (i.e., a linear least-squares discriminant procedure). In normalization, the mean of the data is subtracted from each value and the result is divided by the standard deviation. In the *None* method, no normalization was done. In the *Current* method, the mean and standard deviation were computed from the current session. In the *Last* method, the mean and standard deviation were computed from the immediately preceding session. In the *Recur* method, a recursive estimate of the mean and standard deviation was computed by weighting the previous estimate 0.9 and the data from the immediately preceding trial 0.1.

The None method is a standard nonadaptive method normally used online. The First method is a standard offline analysis method that could not be used online. The Last and Recur methods are adaptive methods that could be used online.

RESULTS

Adaptation of Feature Weights

Figure 1 shows the accuracies for the SMR-based BCI (50% chance accuracy) and the P300-based BCI (2.8% chance accuracy) with the four different methods for determining the feature weights. An analysis of variance with BCI (SMR-based or P300-based) as a between-subjects effect and method (Cross, First, Last, or All) as a within-subjects effect resulted in a significant main effect of BCI ($F = 27.12$, $p < 0.0001$) and a BCI-by-method interaction ($F = 22.24$, $p < 0.0001$).

This significant BCI-by-method interaction was due to the fact that accuracy did not vary with method for the P300-based BCI, but did vary with method for the SMR-based BCI. As Figure 1 shows, SMR-based BCI accuracy was highest for the Cross method, somewhat lower for the adaptive Last and All methods, and lowest for the nonadaptive First method. Post-hoc analyses (Newman-Keuls test) indicated that accuracy for the Cross method was significantly higher than those for the other methods ($p < 0.05$ in all cases). While this method could not be used online, its performance does provide information about the

statistics of the SMR data (see below). Of greatest practical importance, the adaptive Last and All methods produced similar results and were significantly better than the nonadaptive First method ($p < 0.01$ in both cases). This indicates that continually adapting the feature weights during online SMR-based BCI operation improves performance. In contrast, all four methods gave similar accuracy for the P300-based BCI.

Adaptive Normalization

Figure 2 shows the accuracies for the SMR-based BCI (50% chance accuracy) and the P300-based BCI (2.8% chance accuracy) of the four different normalization methods. An analysis of variance with BCI (SMR-based or P300-based) as a between-subjects effect and normalization method (None, Current, Last, Recur) as a within-subject effect resulted in a significant main effect of normalization ($F = 4.29$, $p < 0.0146$) and a significant BCI-by-normalization-method interaction ($F = 11.63$, $p < 0.0001$).

Post-hoc tests indicated that SMR-based BCI accuracy was similar for the Current and Recur methods and that both of these methods were better than the None and Last methods ($p < 0.05$ in all cases). The Current method, unlike the other three, could not be used online. Thus, the important finding is that, for the SMR-based BCI, the adaptive Recur method was better than the nonadaptive None method or the adaptive Last method. In contrast, the accuracies of the four methods did not differ significantly for the P300-based BCI.

DISCUSSION

This study asked whether two different kinds of ongoing adaptation (i.e., adaptation of feature weights and adaptive normalization) can improve the performance of an SMR-based or a P300-based BCI. The results were markedly different for the two BCIs. Consistent with the suggestion of Sellers et al. (2007), the SMR-based BCI benefited from either kind of adaptation, while the P300-based BCI benefited from neither.

Continually adapting feature weights is a form of supervised learning and thus requires knowledge of the user's intention (i.e., the correct outcome). This was accomplished in the present study by instructing the user what to select. In a practical SMR-based BCI system, to be used by people in their daily lives, this could be accomplished with periodic calibration runs. In contrast, continually adapting the normalization procedure is unsupervised, and thus could easily be incorporated into a practical BCI system.

The fact that the two adaptive methods examined here did not improve P300-based BCI accuracy does not necessarily mean that P300 statistics are stationary. What it does appear to indicate is that P300 statistics do not vary over time as SMR statistics do. It remains possible that P300 statistics vary over time due to other factors (e.g., processing load) and thus require different adaptive methodologies. Furthermore, P300 statistics do appear to vary across users. For example, Thulasidas et al. (2006) showed that support vector machine (SVM) models parameterized on the individual user's data performed much better than models parameterized on other users' data. Thus, while there is little evidence to support the use of methods that adapt to systematic changes in P300 statistics over time, it does appear to be advantageous to adapt to user-specific statistics.

The non-stationary nature of the BCI signal provides a rationale for the design of an adaptive BCI system. Vidaurre et al (2006) describe a BCI based on an adaptive classifier. They used a quadratic discriminant analysis with adaptive estimation of the covariance matrix. Vidaurre et al (2006) present some interesting illustrations of how two-dimensional projections of the feature distributions change from one session to the next. Similar results were reported by Shenoy et. al. (2006) who examined several adaptive methods through

offline analysis SMR data. They found the most dramatic difference in statistics between a motor imagery session without feedback and an initial feedback session. They concluded that a small one-time bias adjustment, similar to our adaptive normalization, is worthwhile. Shenoy et al (2006) assert that the mental state of users differs between offline calibration and online feedback sessions, providing a rationale for adaptation. While this is probably true for SMR-based systems it does not appear to hold for current P300 systems (McFarland et al, 2011).

The present study, which examines the impact of adaptation over multiple feedback sessions, indicates that continuing adaptive normalization (i.e., by the Recur rather than the Last method) is beneficial. It is also practical for use in actual BCI operation as it does not require special calibration trials in which users are told what outcome to produce. In contrast to Shenoy et al (2006), we also found that adaptation of the feature weights was beneficial. This difference could be due to our analysis of multiple sessions and/or our use of fewer features, which may reduce the risk of overfitting. This use of fewer features makes ongoing adaptation more practical, because it reduces the need for large amounts of training data. Krausz et al. (2003) report that fast adaptation of parameters during training was not necessary. They suggest that a classifier could be updated at the beginning of each session. These studies illustrate the non-stationary nature of ERD/ERS data. Demonstrating the non-stationary nature of statistics describing the data provides an objective rationale for use of adaptive methods. Unfortunately these studies have been rather short-term to date.

The fact that, in the analysis of feature weight adaptation, the highest SMR-based BCI accuracy was produced by the leave-one-out cross-validation method indicates that the best estimate of the SMR covariance matrix is provided by data from the current session. While this method could not be realized online it does appear to be consistent with the results of Vidaurre et al (2006) who found that continuously updating a classifier produced better classification of SMR data than periodic updates.

The present results were obtained through an off-line analysis of data collected during real-time BCI operation with feedback. Wolpaw et al (2002) discussed several types of adaptation that might occur with a BCI system. These include adaptation to characteristics of individual users, spontaneous variation in signals, and the co-adaptation of user and BCI system. Since offline analysis cannot model user-system co-adaptation there is no guarantee that these offline results will generalize completely to on-line performance. However they do provide useful information, particularly in contrasting SMR and P300 systems, which clearly are different in terms of systematic variation in statistics across sessions. While offline analysis is somewhat limited in generality it is both practical and effective in dealing with some of these issues.

In sum, the present study shows that, for an SMR-based BCI, and for at least the user's initial six sessions, continuing adaptation of feature weights or continuing adaptive normalization can improve performance. In contrast, these adaptive methods do not appear to improve the performance of a P300-based BCI.

Acknowledgments

This work was supported by grants from NIH (HD30146 (NCMRR, NICHD) and EB00856 (NIBIB & NINDS)) and the James S. McDonnell Foundation.

References

- Blankertz B, Muller RK, Krusienski DJ, Schalk G, Wolpaw JR, Schlogl A, Pfurtscheller G, Millan JR, Schroder M, Birbaumer N. The BCI Competition III. Validating alternative approaches to actual BCI problems. *IEEE Trans Neural Syst Rehabil Eng.* 2006; 14(2):153–9. [PubMed: 16792282]
- Fabiani GE, McFarland DJ, Wolpaw JR, Pfurtscheller G. Conversion of EEG activity into cursor movement by a brain-computer interface (BCI). *IEEE Trans Neural Syst Rehabil Eng.* 2004; 12(3): 331–8. [PubMed: 15473195]
- Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing brain potentials. *EEG Clin Neurophysiol.* 1988; 70(6):510–23.
- Ince NF, Arica S, Tewfik A. Classification of single trial motor imagery EEG recordings with subject adapted non-dyadic arbitrary time-frequency tilings. *J Neural Eng.* 2006; 3(3):235–44. [PubMed: 16921207]
- Krausz G, Scherer R, Korisek G, Pfurtscheller G. Critical decision-speed and information transfer in the Graz brain-computer interface. *Appl Psychophysiol Biofeedback.* 2003; 28(3):233–40. [PubMed: 12964454]
- Krusienski DJ, Sellers EW, McFarland DJ, Vaughan TM, Wolpaw JR. Toward enhanced P300 speller performance. *J Neurosci Methods.* 2008; 167(1):15–21. [PubMed: 17822777]
- McFarland DJ, Sarnacki WA, Vaughan TM, Wolpaw R. Brain-computer interface (BCI) operation: signal and noise during early training sessions. *Clin Neurophysiol.* 2005; 116(1):56–62. [PubMed: 15589184]
- McFarland DJ, Anderson CW, Muller K-R, Schlogl A, Krusienski D. BCI Meeting 2005- Workshop on BCI signal processing: Feature extraction and translation. *IEEE Trans Neural Syst Rehabil Eng.* 2006a; 14(2):135–7. [PubMed: 16792278]
- McFarland DJ, Krusienski DJ, Wolpaw JR. Brain-computer interface signal processing at the Wadsworth Center: mu and sensorimotor beta rhythms. *Prog Brain Res.* 2006b; 159:411–9. [PubMed: 17071245]
- McFarland DJ, McCane LM, David SV, Wolpaw JR. Spatial filter selection for EEG-based communication. *Electroencephalogr Clin Neurophysiol.* 1997; 103(3):386–94. [PubMed: 9305287]
- McFarland DJ, Sarnacki WA, Townsend G, Vaughan T, Wolpaw JR. The P300-based brain-computer interface (BCI): Effects of stimulus rate. *Clin Neurophysiol.* 2011; 122(4):731–7. [PubMed: 21067970]
- Schalk G, McFarland DJ, Hinterberger T, Birbaumer N, Wolpaw JR. BCI2000: A general-purpose Brain-Computer interface (BCI) system. *IEEE Trans Biomed Eng.* 2004; 51(6):1034–43. [PubMed: 15188875]
- McFarland DJ, Wolpaw JR. Sensorimotor rhythm-based brain-computer interface (BCI): model order selection for autoregressive spectral analysis. *J Neural Eng.* 2008; 5(2):155–62. [PubMed: 18430974]
- Muller KR, Anderson CW, Birch GE. Linear and non-linear methods for brain-computer interfaces. *IEEE Trans Neural Syst Rehabil Eng.* 2003; 11(2):1659.
- Ramoser H, Wolpaw JR, Pfurtscheller G. EEG-based communication: Evaluation of alternative signal prediction methods. *Biomed Tech.* 1997; 42(9):226–33.
- Sellers EW, Krusienski DJ, McFarland DJ, Vaughan TM, Wolpaw JR. A P300 event-related brain-computer interface (BCI): The effects of matrix size and interstimulus interval on performance. *Biol Psychol.* 2006; 73(3):242–252. [PubMed: 16860920]
- Sellers, EW.; Krusienski, DJ.; McFarland, DJ.; Wolpaw, JR. Noninvasive brain-computer interface research at the Wadsworth Center. In: Dornhege, G.; Millan, J.; Hinterberger, T.; McFarland, D.; Muller, KR., editors. *Towards Brain-Computer Interfacing*. MIT Press; 2007. p. 31–42.
- Sharbrough F, Chatrian CE, Lesser RP, Luders H, Nuwer M, Picton TW. American electroencephalographic society guidelines for standard electrode position nomenclature. *J Clin Neurophysiol.* 1991; 8:200–2. [PubMed: 2050819]
- Shenoy P, Krauledat M, Blankertz B, Rao RPN, Muller KR. Towards adaptive classification for BCI. *J Neural Eng.* 2006; 3(1):13–23.

- Sykacek P, Roberts S, Stokes M. Adaptive BCI based on variational Bayesian Kalman filtering: an empirical evaluation. *IEEE Trans Biomed Eng.* 2004; 51(5):719–27. [PubMed: 15132497]
- Thulasidas M, Guan C, Wu J. Robust classification of EEG signal for brain-computer interface. *IEEE Trans Neural Syst Rehabil Eng.* 2006; 14(1):24–9. [PubMed: 16562628]
- Vidaurre C, Schlogl A, Cabeza R, Scherer R, Pfurtscheller G. A fully on-line adaptive BCI. *IEEE Trans Biomed Eng.* 2006; 53(8):1214–19. [PubMed: 16761852]
- Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. *Clin Neurophysiol.* 2002; 113(6):767–91. [PubMed: 12048038]
- Wolpaw JR, McFarland DJ, Neat GW, Forneris CA. An EEG-based brain-computer interface for cursor control. *Electroencephalogr Clin Neurophysiol.* 1991; 78(3):252–259. [PubMed: 1707798]
- Wolpaw JR, McFarland DJ. Multichannel EEG-based brain-computer communication. *Electroencephalogr Clin Neurophysiol.* 1994; 90(6):444–9. [PubMed: 7515787]
- Wolpaw JR, McFarland DJ. Control of a two-dimensional movement signal by a non-invasive brain-computer interface in humans. *Proc Natl Acad Sci USA.* 2004; 101(51):17849–54. [PubMed: 15585584]

Figure 1. Three approaches to BCI design. The arrows indicate whether the BCI, the user, or both adapt to optimize and maintain BCI performance. (Adapted from McFarland et al., 2006).

Figure 2.

Accuracies for the SMR-based BCI (50% chance accuracy) (black) and the P300-based BCI (2.8% chance accuracy) (gray) with the four different methods for determining the feature weights. Error bars indicate SEM. Data were not normalized. For the SMR-based BCI, the Cross method (which could not be used online) was significantly better than the adaptive Last and All methods, which were significantly better than the nonadaptive First method. For the P300-based BCI, the four methods performed similarly. See text for details of methods.

Figure 3.

Accuracies for the SMR-based BCI (50% chance accuracy) and the P300-based BCI (2.8% chance accuracy) of the four different normalization methods. Error bars indicate SEM. There was no adaptation. For the SMR-based BCI, the Current method (which could not be used online) and the adaptive Recur method performed similarly and significantly better than the nonadaptive None method or the adaptive Last method. For the P300-based BCI, the four methods did not differ significantly in accuracy. See text for details of methods.