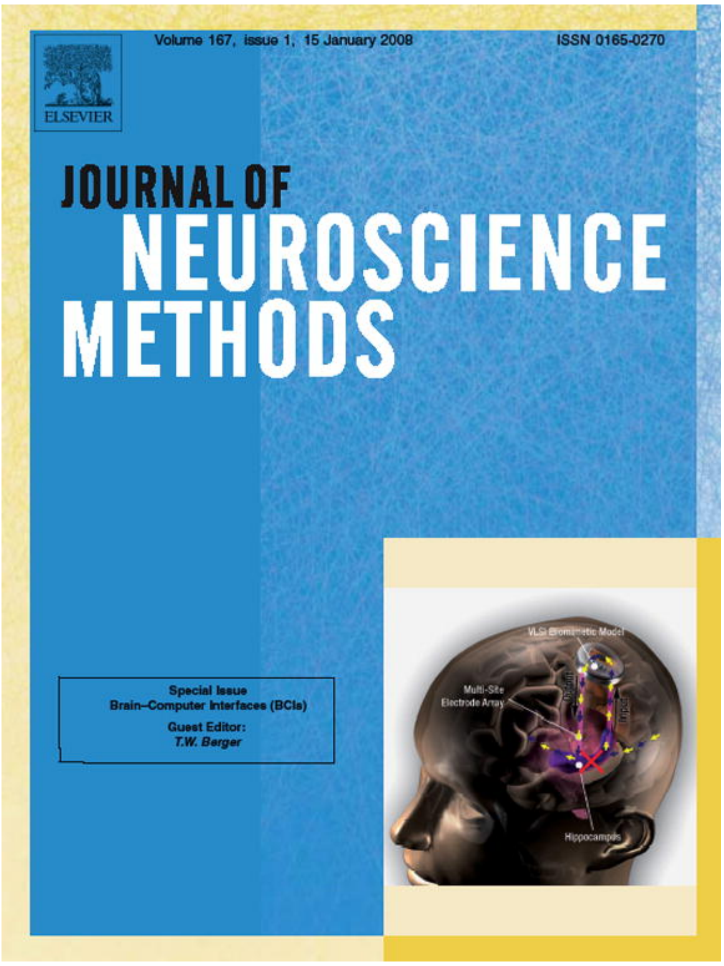
Volume 167, issue 1, 15 January 2008
ISSN 0165-0270
ELSEVIER
JOURNAL OF
NEUROSCIENCE
METHODS
Special Issue
Brain–Computer Interfaces (BCIs)
Guest Editor:
T.W. Berger
VLSI Biomimetic Model
Multi-Site
Electrode Array
Hippocampus

ELSEVIER

Journal of Neuroscience Methods 167 (2008) 15–21

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Toward enhanced P300 speller performance

D.J. Krusienski [a,*], E.W. Sellers [b], D.J. McFarland [b], T.M. Vaughan [b], J.R. Wolpaw [b]

[a] *University of North Florida, Division of Engineering, 4567 St. Johns Bluff Road, South Jacksonville, FL 32224-2645, United States*
[b] *Wadsworth Center, New York State Department of Health, E1000 Empire State Plaza, Albany, NY 12201, United States*

Received 3 April 2007; received in revised form 18 July 2007; accepted 18 July 2007

Abstract

This study examines the effects of expanding the classical P300 feature space on the classification performance of data collected from a P300 speller paradigm [Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroenceph Clin Neurophysiol 1988;70:510–23]. Using stepwise linear discriminant analysis (SWLDA) to construct a classifier, the effects of spatial channel selection, channel referencing, data decimation, and maximum number of model features are compared with the intent of establishing a baseline not only for the SWLDA classifier, but for related P300 speller classification methods in general. By supplementing the classical P300 recording locations with posterior locations, online classification performance of P300 speller responses can be significantly improved using SWLDA and the favorable parameters derived from the offline comparative analysis.

Keywords: Brain–computer interface; Event related potentials; P300 speller; Stepwise linear discriminant analysis

1. Introduction

A brain–computer interface (BCI) is a device that uses brain signals to provide a non-muscular communication channel (Wolpaw et al., 2002), particularly for individuals with severe neuromuscular disabilities. The P300-event related potential is an evoked response to an external stimulus that is observed in scalp-recorded electroencephalography (EEG). The P300 response has proven to be a reliable signal for controlling a BCI (Farwell and Donchin, 1988). Farwell and Donchin (1988) describe the P300 speller, which presents a selection of characters arranged in a 6×6 matrix. The user focuses attention on one of the 36 character cells of the matrix while each row and column of the matrix is intensified in a random sequence. The row and column intensifications that intersect at the attended cell represent the target stimuli, which occur with a probability of 1/6. The rare presentation of the target stimuli in the random sequence of stimuli constitutes an Oddball Paradigm (Fabiani et al., 1987) and will elicit a P300 response to the target stimuli. With proper P300 feature selection and classification, the attended character of the matrix can be identified and communicated.

A variety of feature extraction and classification procedures such as stepwise linear discriminate analysis (SWLDA) (Donchin et al., 2000; Sellers and Donchin, 2006), wavelets (Bostanov, 2004), support vector machines (Kaper et al., 2004; Meinicke et al., 2002; Thulasidas et al., 2006), and matched filtering (Serby et al., 2005) have been implemented, improving the performance beyond that originally reported in (Farwell and Donchin, 1988). Based on multiple studies in healthy volunteers (Donchin et al., 2000; Sellers and Donchin, 2006; Serby et al., 2005), and initial studies in persons with physical disability (Vaughan et al., 2006), the P300 speller has potential to serve as an effective communication device for persons who have lost or are losing the ability to write and speak. An individual with advanced-stage ALS reported the P300 speller to be superior and preferential to his modern eye-gaze system and uses the BCI 4–6 h/day for e-mail and other computer applications (Vaughan et al., 2006). Initial reports from other disabled individuals currently testing the P300 speller BCI also indicate that its speed, accuracy, and ease of use are superior or competitive with other assistive technologies.

Up to the present, BCI-related P300 research has focused almost exclusively on signals from standard P300 scalp locations (i.e., Fz, Cz, Pz). While recent offline evaluations suggest that the use of additional locations, particularly posterior sites, may improve classification accuracy (BCI, 2003; BCI, 2005; Blankertz et al., 2004, 2006; Kaper et al., 2004; Spencer et al., 2001; Vaughan et al., 2003), this possibility has not been formally addressed in comprehensive offline and online studies.

* Corresponding author. Tel.: +1 904 620 1343; fax: +1 904 620 1391.
E-mail address: deankrusienski@ieee.org (D.J. Krusienski).

0165-0270/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2007.07.017

To address this possibility, the present study explores the value of incorporating information from electrode locations that are not traditionally associated with the P300 response. In addition, several data preprocessing and model parameters are evaluated to assess the relative effects with respect to the new spatial information. Using a SWLDA classifier, both offline and online results obtained from 64-channel data show that some of the most discriminable EEG features evoked by the P300 speller occur at posterior electrodes (namely PO_7, PO_8, Oz), and that these features can significantly improve classification performance when used in conjunction with the classical P300 feature space (i.e., EEG features at electrodes Fz, Cz, Pz (Sharbrough et al., 1991)). The results are also relevant to current speculations concerning the nature of the neural processes underlying P300-based BCI operation and the question of their dependence on gaze direction.

2. Data collection

2.1. Participants

Seven able-bodied people were the participants in this study. The demographics and previous speller matrix experience of the participants are listed in Table 1. The participants varied in their previous BCI experience, but all participants had either no experience or relatively few sessions with a P300-based BCI system. The study was approved by the New York State Department of Health Institutional Review Board, and each participant gave informed consent.

2.2. Task, procedure, and design

The participant sat upright in front of a video monitor and viewed the matrix display. The task was to focus attention on a specified letter of the matrix and silently count the number of times the target character intensified, until a new character was specified for selection. All data was collected in the copy speller mode: words were presented on the top left of the video monitor and the character currently specified for selection was listed in parentheses at the end of the letter string (see Fig. 1). Each session consisted of nine experimental runs; each run was composed of a word or series of characters chosen by the investigator. This set of characters spanned the set of characters contained in the matrix and was consistent for each participant and session. The rows and columns were intensified for 100 ms with 75 ms

Table 1
Participant demographics

	Age	Gender	No. of prior sessions
Participant A	24	Male	6
Participant B	47	Male	0
Participant C	50	Female	12
Participant D	36	Male	6
Participant E	27	Male	0
Participant F	24	Male	0
Participant G	32	Male	0

Fig. 1. The 6 × 6 matrix used in the current study. A row or column intensifies for 100 ms every 175 ms. The letter in parentheses at the top of the window is the current target letter "D." A P300 should be elicited when the fourth column or first row is intensified. After the intensification sequence for a character epoch, the result is classified and online feedback is provided directly below the character to be copied.

between intensifications. One complete cycle of six row and six column intensifications constitutes a sequence and 15 sequences constitutes a character epoch. Specifically, the classification was performed after every row and column has been intensified 15 times. Each session consisted of 36 character epochs, equivalent to 6480 stimuli (row/column intensifications). A single session, lasting approximately 1 h, per participant was collected per day. Five sessions, collected over a period of weeks, were obtained from each of the seven participants. Two additional sessions were collected from five of the participants for verification of the offline analysis.

For each channel used in the analysis, 800-ms segments of data (192 samples) following each intensification were extracted for the offline analysis. The data segments were concatenated by channel for each intensification, creating a single feature vector corresponding to each stimulus. The N features × 6480 stimuli observation matrix was used to derive the SWLDA weights for each participant, where N = [# channels × 192 samples].

2.3. Data acquisition

The EEG was recorded using a cap (Electro-Cap International Inc.) embedded with 64 electrode locations distributed over the entire scalp, based on the International 10–20 system (Sharbrough et al., 1991). All 64 channels were referenced to the right earlobe, and grounded to the right mastoid. The EEG was bandpass filtered 0.1–60 Hz and amplified with a SA Electronics amplifier (20,000×), digitized at a rate of 240 Hz, and stored. All aspects of data collection and experimental procedure were controlled by the BCI2000 system (Schalk et al., 2004).

3. Stepwise linear discriminant analysis

Determining the presence or absence of a P300 evoked potential from EEG features can be considered a binary classification problem with a decision hyper-plane defined by:

$$w \cdot x - b = 0 \tag{1}$$

where x is the feature vector as described in Section 2.2, w a vector of feature weights, and b is the bias term. However, because it is assumed that a P300 is elicited for one of the six row/column intensifications, and that the initial results indicate that the P300 response is invariant to row/column stimuli, the resultant classification is taken as the maximum of the sum of scored feature vectors for the respective rows, as well as for the columns:

$$\text{predicted row} = \max_{\text{rows}} \left[\sum_{i_{\text{row}}} w \cdot x_{i_{\text{row}}} \right] \tag{2}$$

$$\text{predicted column} = \max_{\text{columns}} \left[\sum_{i_{\text{column}}} w \cdot x_{i_{\text{column}}} \right] \tag{3}$$

This design selects the response with the largest positive distance from the trained separating hyper-plane, which is ideally analogous to selecting the response that strongly represents the characteristic P300 as defined by the training data. The predicted character is located at the intersection of the predicted row and column in the matrix.

Stepwise linear discriminant analysis (Draper and Smith, 1981) is a technique for selecting suitable predictor variables to be included in a multiple regression model as given in Eq. (1). For binary classification tasks such as this, the linear discriminant and least-squares regression solutions are equivalent. A combination of forward and backward stepwise regression is implemented. Starting with no initial model terms, the most statistically significant predictor variable having a p-value <0.1, is added to the model. After each new entry to the model, a backward stepwise regression is performed to remove the least significant variables, having p-values >0.15. This process is repeated until the model includes a predetermined number of terms, or until no additional terms satisfy the entry/removal criteria.

The SWLDA algorithm can be considered efficient because the terminating heuristic is implemented in such a way that suitable features are selected in a non-exhaustive manner. The only required parameters, the maximum model order and the termination heuristic, are intuitive and can be easily gauged based on the expected characteristics of the data. In a sense, SWLDA has the advantage of having automatic feature extraction. Because insignificant terms are removed from the model (i.e., weights are set to zero), using less training data is less likely to corrupt the classification result because insignificant features are completely eliminated from the model. Though SWLDA can be tuned to provide faster convergence by limiting the model order or termination heuristic, it is not guaranteed to be convergent and will not provide a model if the heuristic cannot be satisfied. However, this typically occurs only if the model is inadequate or if there is not discriminable information contained within the features. When properly configured, this result can be used to conclude that P300 evoked potentials are not present in the session.

4. Analysis protocol

In the previous work on SWLDA for classifying P300 responses (BCI, 2003; Farwell and Donchin, 1988; Sellers and Donchin, 2006), only channels Fz, Cz, and Pz were used for analysis. However, the posterior response seems to provide significant additional discriminative information for the P300 speller (BCI, 2003; BCI, 2005; Blankertz et al., 2004, 2006; Kaper et al., 2004; Spencer et al., 2001; Vaughan et al., 2003). Thus far, neither the temporal attributes of this posterior response nor its relationship to the central P300 response have been characterized. Because of this, the present study examines several aspects of the feature space in order to determine if the classification performance can indeed be improved by incorporating additional channels and possibly altering the data preprocessing. The effects of the following four factors on SWLDA classification of offline P300 speller data are evaluated: channel set, reference, decimation factor, and maximum features. A description of each of these factors is given below.

4.1. Channel set

Several overlapping and non-overlapping subsets of channels are examined to compare the relative emphasis of spatial information on classification as well as to define a robust, minimum set that can serve as a starting point for future analysis. The four individual channel sets selected for analysis are illustrated in Fig. 2. Channel set 1 represents the classical channels used for extracting the P300 response. Channel set 2 represents the posterior regions that have strong correlations with desired matrix targets. Channel set 3 is the union of channel sets 1 and

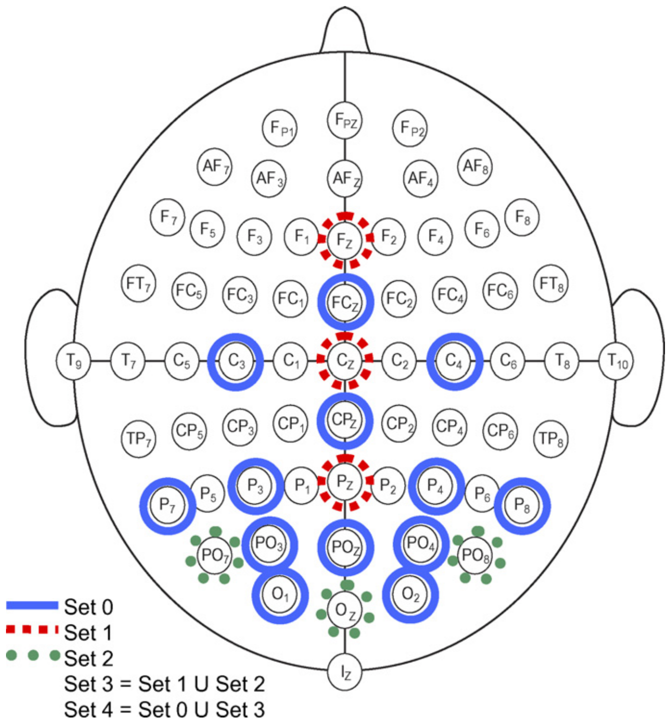

Fig. 2. The 64-channel electrode montage and the channel sets. Set 0 is a subset defined purely for illustration purposes, sets 1–4 were used in the analysis.

2, to demonstrate the interaction effects of the central and posterior information. Channel set 4 represents an expanded number of channels from the key regions to demonstrate the effects of providing SWLDA additional spatial diversity for selection. Channel set 0 is defined purely for illustration purposes as the residual subset of channel set 4 not provided by the other channel sets.

4.2. Reference

To compare whether the global effects interact with the localized channel sets, the ear referenced data is compared to data using a common average reference (CAR) of all 64 channels.

4.3. Decimation factor

Because of the high sampling rate of the recordings relative to the low frequency of the P300 response, a dimensionality reduction for removal of redundant features is beneficial for classification. Rather than simply decimating the data, the data are segmented into blocks having length equal to the selected decimation factor. The mean of these blocks is calculated and used as the feature, effectively smoothing and decimating the data. This is equivalent to passing the data through a moving average filter before decimating. Three different decimation factors are examined: 6, 12, and 24, corresponding to sample rates of 40, 20, and 10 Hz, respectively. The basis of the P300 response is believed to lie within these frequency ranges (Intriligator and Polich, 1994).

4.4. Maximum features

The maximum number of features permitted for inclusion in the regression model is evaluated to determine a reasonable value for a given number of spatial/temporal features that demonstrate good generalizability to test data. Setting the number of maximum features too low may exclude relevant classification information, whereas values too high may begin to over-fit the data.

The classification performance of each combination of factors was evaluated using the offline data. The total number of features available to SWLDA for each combination is given in Table 2. For each of the seven participants and combinations of factors, SWLDA was used to derive feature weights from the participant's first session data only, using all 15 intensification sequences. These weights were then tested on all four subsequent sessions. For the test sessions, the feature vectors for each subsequent intensification in the sequence (up to 15) were averaged by corresponding row/columns for each character epoch and classified to compare performance for a minimum number of intensifications. The results presented are the averages of the resulting performances from each of the four test sessions.

Table 2
Total features available for a given channel set and decimation factor

No. of electrodes	Decimation factor		
	24	12	6
3 (sets 1 and 2)	24	48	152
6 (set 3)	48	96	304
19 (set 4)	96	192	608

5. Results

5.1. Offline evaluation

The four factors of primary interest were examined separately offline. Fig. 3 illustrates the fundamental effects of channel set, reference, decimation, and maximum features, respectively. Because all combinations of factors were initially evaluated, factors that are clearly superior or show consistent performance across conditions are fixed for presentation purposes. For example, the factors of decimation and maximum features do not result in significantly different performance for any of the conditions, but values of 12 for decimation and 60 for maximum features appear to be superior for practically all conditions and are therefore fixed to those values for the analysis of the other factors. The ear reference is used to compare the channel sets because the CAR incorporates global spatial information, which could possibly mask the true contribution of the smaller, localized channel sets. The ear reference is also used in the decimation factor plot because there is not an appreciable difference between it and the CAR. The best reference for each participant is compared for the maximum features factor because it demonstrates more directly the effect of the maximum features in a best-case scenario. The best channel set was also selected across factors because the results are consistent within participants and more representative of how the factors affect the best-case scenario. Table 3 provides a summary of the factors used in the plots.

A repeated measures analysis of variance (ANOVA) was performed on the four factors. The only factor that yielded a significant effect was channel set ($F(3, 24) = 5.02$, $p = 0.0077$). As shown in Fig. 3, channel sets 3 and 4 performed significantly better than channel sets 1 and 2. No difference was observed between sets 3 and 4, and no difference was observed between sets 1 and 2 (Fisher LSD = 18.44, $p < 0.05$).

Overall, the CAR is not different from the ear reference for the best channel set ($F(1, 24) = 0.01$, $p = 0.9064$). This indicates that the majority of essential P300 response information is contained in channel sets 3 and 4, and that the CAR contributes little or no additional information to these two specific channel sets.

Table 3
Factors used in Fig. 3

	Ch. set	Ref.	Dec. fact.	Max. feat.
Ch. set	X	Ear	12	60
Ref.	Best	X	12	60
Dec. fact.	Best	Ear	X	60
Max. feat.	Best	Best	12	X

Each row indicates a single dependent factor considered in Fig. 3. The columns give the fixed independent factor settings used for the analysis of each dependent factor. "Best" indicates the value of the factor that gives the highest overall performance for a particular participant, consistent across all conditions.

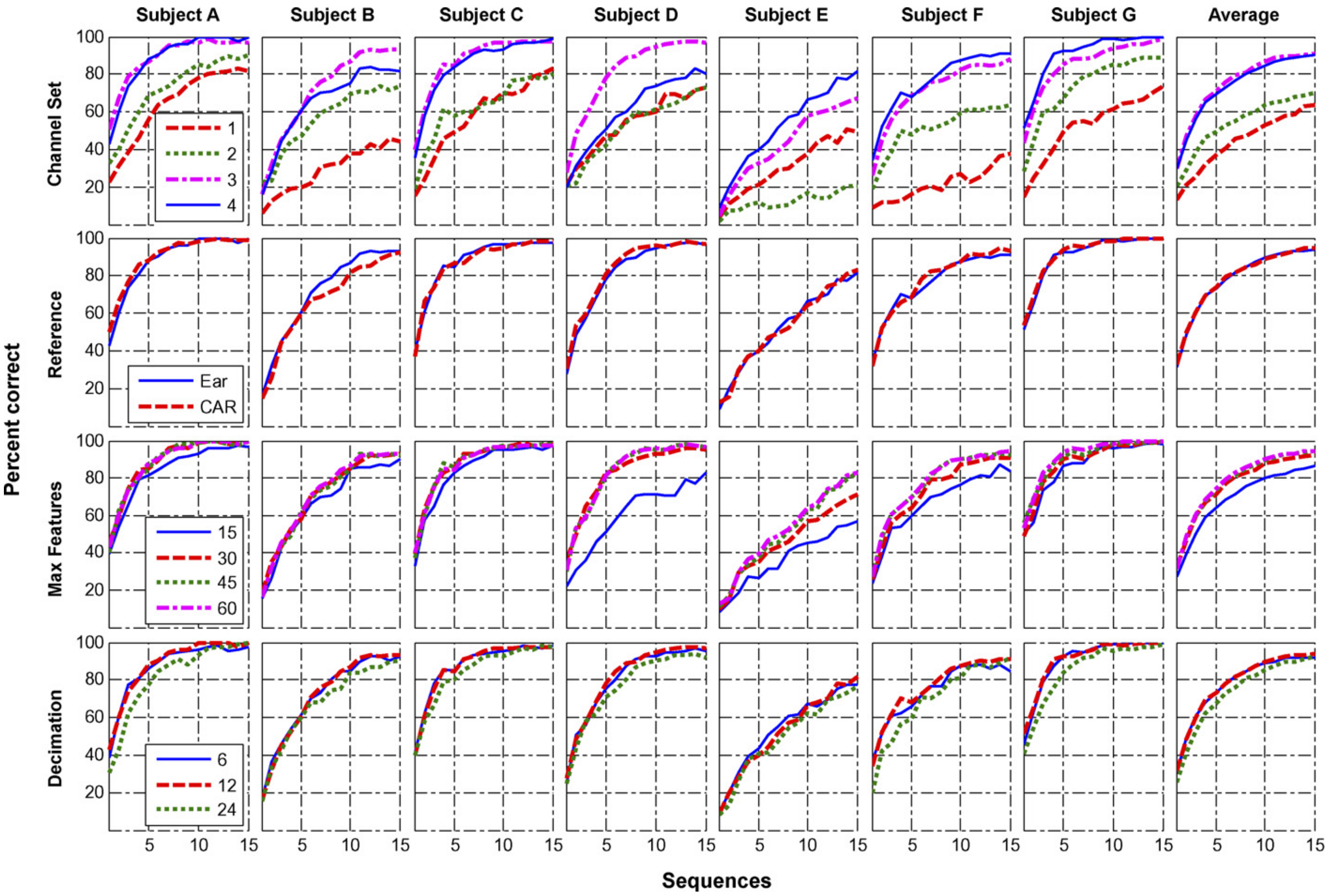

Fig. 3. Performance curves at each level of the four factors for all 15 stimulus intensification sequences. Each of the four rows of plots indicates a single factor for each participant. The rightmost column of plots shows the average for each factor across participants.

The factor of decimation did not yield a significant difference for the tested values of 6, 12, and 24 ($F(2, 24) = 0.17$, $p = 0.8459$). Despite the lack of statistical significance, the results consistently show that a decimation factor of 24 is less desirable than 6 or 12. It is possible that the current analysis lacked sufficient power to detect a significant difference.

The statistical analysis performed on the factor of maximum features also resulted in no statistical differences between the four levels ($F(3, 24) = 1.09$, $p = 0.373$). A consistently lower performance for a maximum feature set of 15 is observed but is not identified as significant from the statistical analysis. Again, this could be due to a lack of sufficient power to detect a significant difference.

5.2. Online performance

Following the offline analysis, an additional set of two experimental sessions was performed for five of the seven participants. The additional sessions served two purposes. First, to validate the offline results in an online situation. Second, to examine differences between the commonly used midline electrodes (set 1) and the set of electrodes that classifies most accurately for each

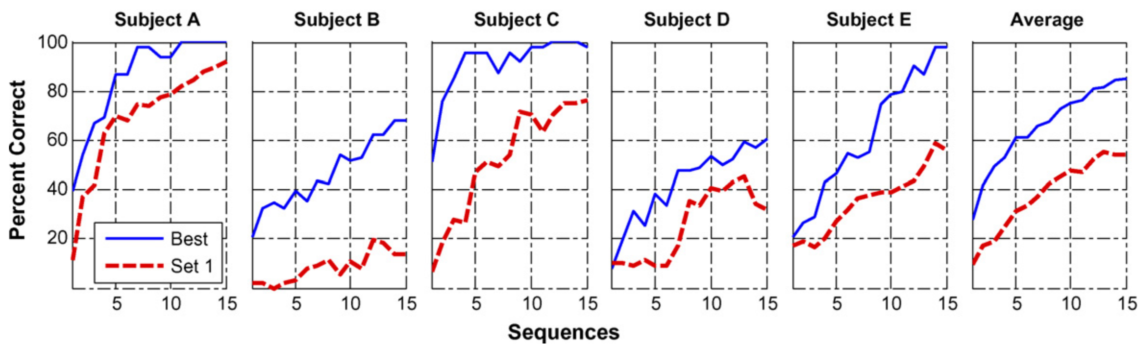

Fig. 4. Online Performance for each participant using the best set of feature weights (solid line) or set 1 weights (Fz, Cz, and Pz only) (dashed line). Each participant completed two sessions. Half of each session was conducted with the best weights and half with set 1 weights, in counterbalanced order.

participant. Based on the offline findings, two sets of classification weights were derived for online use. One set used only the midline electrodes and a second set was the set that performed best for each of the five participants. The factors of reference, decimation, and maximum iterations were set to ear reference, 12, and 60, respectively. Each session was conducted using both sets of weights to provide online feedback. Half of each session was conducted using the SWLDA weights representing the best channel set for the particular participant, the other half was conducted using channel set 1. The application of the different weights used for each session was counterbalanced. The elapsed time between the initial weight generating session and the online verification sessions was 11 months for Participant E, 7 months for Participants B and C, and 2 months for Participants A and D.

The results are shown in Fig. 4. The figure shows that all five participants performed best with the set of weights derived from the large electrode set. In addition, the expanded set of electrodes performed approximately 25% better, on average.

6. Discussion

It is evident from the results that an essential factor in optimizing the performance of the P300 classifier is the identification of an appropriate channel set for the individual. Channel set was the only factor that yielded a significant statistical effect on classification rates; however, the analyses of the other factors have provided important information. The analyses conducted here suggest that, on average, a decimation of 12, ear reference, maximum features of at least 45, and either channel set 3 or 4 provide the best classification parameters for SWLDA.

In general, the performance of the classical P300 electrode set (set 1) was inferior to the expanded sets of electrodes (sets 3 and 4), and statistically equivalent to the posterior set (set 2), indicating that the posterior electrodes contribute additional information to P300 speller classification. Performance is significantly increased when the central and posterior channel sets are combined in set 3. Moreover, an additional 13 electrodes (set 4) does not produce a significant increase in performance over the performance achieved by the six electrodes in set 3, although set 3 is a subset of set 4. This suggests that including 19 electrodes results at best in no appreciable improvement from six electrodes, and at worst in a small amount of overfitting. Thus, SWLDA is not exactly optimal in terms of feature extraction because introducing additional input features to the model that provide little additional information can lead to a decrease in performance in some instances. Based on the results of this study, channel sets 3 and 4 serve as reasonable starting points that can be minimally extended or pruned to identify a suitable channel set for a particular participant. It should be noted that SWLDA actually performs pruning since feature from all input channels are not guaranteed to be weighted by this procedure. Alternative feature extraction methods in a BCI context are given in Lal et al. (2004) and Mueller et al. (2004). Additional evidence for the efficacy of using electrodes that include posterior, central, and frontal electrodes is provided in Spencer et al. (2001). In this study, PCA was performed on Oddball data and demonstrated that a component extending from Cz to Oz and P7 to P8 accounted for the largest portion of variance, and a component extending from Cz to Fz accounted for the second largest portion of variance.

With the feature weights derived by SWLDA, as described in the offline analysis, online performance confirmed what the offline analyses suggested. That is, performance was superior using expanded electrode sets as compared to the three midline electrodes. In addition, using the expanded set of electrodes, all participants reached accuracy levels of at least 60% correct. Three of the five participants performed above 90% correct with fewer than 15 sequences. This indicates that the classification can be performed on a minimal number of sequences without compromising accuracy, effectively increasing the communication rate. These results demonstrate the stable and robust nature of the EEG response to the P300 speller, because the weights used online were derived from data collected up to 1 year prior to the online sessions. This finding is consistent with results reported by Sellers and Donchin (2006). Interestingly, Participant E's performance was markedly better during the online experiment that occurred nearly a year after the session from which the classification weights were generated. Participant E's performance did not plateau in the early sessions, indicating that the performance could have continued to increase with additional trials. However, based on the improved online performance, the performance during the early sessions could be attributed to the novelty of the task or the nature of the feedback during these sessions that was based on suboptimal classification weights. The drastic decrease in online performance for Participants B and D is presumed to be due to concentration and attentional issues reported by the participants during the online sessions, rather than to obsolete feature weights. However, this is merely speculation and the participants were not systematically or explicitly asked to report on fatigue, attention, concentration, etc., during the sessions.

The relative contribution of the posterior responses raises the possibility that transient ocular responses could be involved and that foveating the target may be a factor in P300 speller performance. This would be important since BCI systems that are dependent upon peripheral movements are perhaps less useful for those with severe motor disabilities (Wolpaw et al., 2002). However, it is by no means certain that P300 speller performance is dependent in this way. First of all, the P300 response can be observed when spatial attention is not a factor, as with the use of auditory stimuli or visual stimuli that appear sequentially in the same location (Sellers and Donchin, 2006). Secondly, spatial attention involves both overt eye gaze and covert processes. Covert visuospatial attention has been shown to modulate neuronal activity (Thompson et al., 2005) and ERP components (Nobre et al., 2000). Thus, definitive information on the relative contribution of overt eye movements to the P300 speller performance will require either studies controlling for the effect of eye movements or studies with individuals lacking voluntary control of eye movements. It is hypothesized that the posterior features contribute to the classification because the visual evoked responses observed at these sites differ depending on whether the target flash is in the focal or peripheral field. An initial inspection of the posterior target responses (focal field) shows that all participants had a prominent event-related negativity between 200 and 250 ms at PO_7, PO_8, and/or Oz (discussed in Sellers et al., 2007). The posterior standard responses (peripheral field) tend to exhibit a sinusoidal steady-state visual evoked pattern at the flashing frequency as expected.

In conclusion, with the addition of posterior features and a SWLDA trained classifier, the participants of this study were able to achieve speed/accuracy levels suitable for practical online communication. When blind tested on Data Set II (P300 Speller) from the 2005 BCI Competition III (BCI, 2005; Blankertz et al., 2006), the SWLDA method (using ear referencing, electrode set 3 in addition to electrodes P_3 and P_4, a decimation factor of 12, and a maximum of 60 model features) classified the Data Set at 92.5% accuracy, results that rank second[1] (to Rakotomamonjy (BCI, 2005)) amidst contributions that primarily employ advanced support vector machine and clustering techniques. Additionally, for each subject, the competition winner performed an extensive optimization of the channel selection using all 64 channels, as opposed to the eight standardized channels used in the proposed SWLDA method. These performance results indicate that a comparatively straightforward and intuitive linear classifier with minimal preprocessing is as viable and capable as many complex, modern classification techniques for the P300 speller. When trained on a feature space expanded to include information from posterior as well as central electrodes, SWLDA's comparatively straightforward and minimal implementation and training requirements, rapid convergence, modeling robustness, and, most importantly, classification accuracy, make it a very appealing and effective approach for practical application of the P300 speller.

Acknowledgements

This work was supported in part by the National Institutes of Health under Grant NICHD HD30146 and Grant NIBIB/NINDS EB00856 and in part by the James S. McDonnell Foundation.

References

BCI Competition II Data Set IIb. Results (Blankertz and Curio); (2003) http://ida.first.fraunhofer.de/projects/bci/competition_ii/results/.

BCI Competition III. Data Set II Results; (2005) http://ida.first.fraunhofer.de/projects/bci/competition_iii/results/.

Blankertz B, Müller KR, Curio G, Vaughan TM, Schalk G, Wolpaw JR, et al. The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials. IEEE Trans Biomed Eng 2004;51(6): 1044–51.

Blankertz B, Müller KR, Krusienski DJ, Schalk G, Wolpaw JR, Schlögl A, et al. The BCI competition III: validating alternative approaches to actual BCI problems. IEEE Trans Neural Syst Rehabil Eng 2006;14(2): 153–9.

Bostanov V. BCI competition 2003-data sets Ib and IIb: feature extraction from event-related brain potentials with the continuous wavelet transform and the *t*-value scalogram. IEEE Trans Biomed Eng 2004;51:1057–61.

Donchin E, Spencer KM, Wijesinghe R. The mental prosthesis: assessing the speed of a P300-based brain–computer interface. IEEE Trans Rehabil Eng 2000;8:174–9.

Draper N, Smith H. Applied regression analysis. 2nd ed. John Wiley and Sons; 1981. pp. 307–12.

Fabiani M, Gratton G, Karis D, Donchin E. Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. Adv Psychophysiol 1987;2:1–78.

Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroenceph Clin Neurophysiol 1988;70:510–23.

Intriligator J, Polich J. On the relationship between background EEG and the P300 event-related potential. Biol Psychol 1994;37:207–18.

Kaper M, Meinicke P, Grossekathoefer U, Lingner T, Ritter H. BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm. IEEE Trans Biomed Eng 2004;51:1073–6.

Lal TN, Schroder M, Hinterberger T, Weston J, Bogdan M, Birbaumer N, et al. Support vector channel selection in BCI. IEEE Trans Biomed Eng 2004;51(6):1003–10.

Meinicke P, Kaper M, Hoppe F, Huemann M, Ritter H. Improving transfer rates in brain computer interface: a case study. NIPS 2002:1107–14.

Mueller KR, Krauledat M, Dornhege G, Curio G, Blankertz B. Machine learning techniques for brain–computer interfaces. Biomed Tech 2004;49(1):11–22.

Nobre AC, Sebestyen GN, Miniussi C. The dynamics of shifting visuospatial attention revealed by event-related potentials. Neuropsychologia 2000;38:964–74.

Schalk G, McFarland DJ, Hinterberger T, Birbaumer N, Wolpaw JR. BCI2000: a general-purpose brain–computer interface (BCI) system. IEEE Trans Biomed Eng 2004;51:1034–43.

Sellers EW, Krusienski DJ, McFarland DJ, Wolpaw JR. Non-invasive brain–computer interface research at the Wadsworth Center. In: Dornhege, Millan, Hinterberger, McFarland, Müller (Eds.). Towards brain–computer interfacing. MIT Press; 2007. pp. 31–42.

Sellers EW, Donchin E. A P300-based brain–computer interface: initial tests by ALS patients. Clin Neurophysiol 2006;117:538–48.

Serby H, Yom-Tov E, Inbar GF. An improved P300-based brain–computer interface. IEEE Trans Neural Syst Rehabil Eng 2005;13:89–98.

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.

Spencer KM, Dien J, Donchin E. Spatiotemporal analysis of the late ERP responses to deviant stimuli. Psychophysiology 2001;38:343–58.

Thompson KG, Biscoe KL, Sata TR. Neuronal basis of covert spatial attention in the frontal eye fields. J Neurosci 2005;25:9479–87.

Thulasidas M, Cuntai G, Wu J. Robust classification of EEG signal for brain–computer interface. IEEE Trans Neural Syst Rehabil Eng 2006;14(1):24–9.

Vaughan TM, McFarland DJ, Schalk G, Sellers E, Wolpaw JR. Multichannel data from a brain–computer interface (BCI) speller using a P300 (i.e., oddball) protocol. Soc Neurosci Abs 2003.

Vaughan TM, McFarland DJ, Schalk G, Sarnacki WA, Krusienski DJ, Sellers EW, et al. The Wadsworth BCI research and development program: at home with BCI. IEEE Trans Neural Sys Rehabil Eng 2006;14(2): 229–33.

Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain–computer interfaces for communication and control. Clin Neurophysiol 2002;113:767–91.

[1] These results were obtained in strict accordance with the BCI Competition III protocol. However, due to a potential conflict of interest, the results were not officially submitted to the competition and neither the results nor the ranking should be regarded as part of the competition.