A Novel Method for Single-Trial Classification in the Face of Temporal Variability

Amar R. Marathe, Anthony J. Ries, and Kaleb McDowell

Human Research and Engineering Directorate, US Army Research Laboratory, Aberdeen Proving Ground, MD 21005, USA
amar.marathe@case.edu, anthony.j.ries2@mail.mil, kgm8@cornell.edu

Abstract. Machine learning techniques have been used to classify patterns of neural data obtained from electroencephalography (EEG) to increase human-system performance. This classification approach works well in controlled laboratory settings since many of the machine learning techniques used often rely on consistent neural responses and behavioral performance over time. Moving to more dynamic, unconstrained environments, however, introduces temporal variability in the neural response resulting in sub-optimal classification performance. This study describes a novel classification method that accounts for temporal variability in the neural response to increase classification performance. Specifically, using sliding windows in hierarchical discriminant component analysis (HDCA), we demonstrate a decrease in classification error by over 50% when compared to other state-of-the-art classification methods.

Keywords: Brain-Computer Interface (BCI), Rapid Serial Visual Presentation (RSVP), Electroencephalography (EEG), HDCA, Sliding HDCA, Temporal Variability, Single-trial, Real-world environment.

1 Introduction

Systems incorporating neural activity using EEG typically use machine learning techniques to classify or predict the occurrence of an action or event. To be useful, these systems must be able to function outside of the controlled confines of a laboratory setting. Moving into more dynamic environments introduces changes in the processing demands of the user as well as uncontrolled variability into the system. Variability of the EEG signal is influenced by an interaction of endogenous processes related to a user's state (e.g. fatigue), exogenous factors related to stimulus properties [1–4], and other system related factors. For example, it has been shown that the latency of the P300 event related potential (ERP) brain response is correlated with stimulus evaluation and reaction time [5, 6]. Stimuli that are easier to categorize produce faster reaction times and earlier P300 peak latencies than those that are more difficult to categorize. Thus, situations where the difficulty of stimulus categorization varies from trial to trial will produce a temporally variable neural response. Optimal performance

D.D. Schmorrow and C.M. Fidopiastis (Eds.): AC/HCII 2013, LNAI 8027, pp. 345–352, 2013.

of systems interpreting neural data must account for the existence of trial by trial temporal variability in the neural response.

Existing methods for single-trial classification can be divided into several categories. Some algorithms operate directly on the multi-channel EEG signals [7–11], while others apply spatial filters to transform the multi-channel EEG signal into a new signal that contains more task-relevant information prior to applying a standard machine-learning classifier [12–21]. Each of these existing methods have been shown to perform well in a specific task; however none of the previous studies has focused on testing the effects of temporal variability on classification performance. In this study, participants performed a rapid serial visual presentation (RSVP) target detection task. ERP analysis shows that the neural data contains large amounts of temporal variability. We show that a novel classification method that accounts for temporal variability can reduce classification error by over 50%.

2 Methods

2.1 Participants

Fifteen participants (9 male, age range 18-57, average age 39.5) volunteered for the current study. Participants provided written informed consent, reported normal or corrected-to-normal vision and reported no history of neurological problems. Fourteen of the fifteen participants were right-handed.

The voluntary, fully informed consent of the persons used in this research was obtained as required by Title 32, Part 219 of the Code of Federal Regulations and Army Regulations 70-25. The investigator has adhered to the policies for the protection of human subjects as prescribed in AR 70-25.

2.2 Stimuli and Procedure

Short video clips were used in a rapid serial visual presentation (RSVP) paradigm [22, 23]. Video clips either contained people or vehicles on background scenes, or only background scenes. Observers were instructed to make a manual button press with their dominant hand when they detected a person or vehicle (targets), and to abstain from responding when a background scene (distractor) was presented. Video clips consisted of five consecutive images each 100ms in duration; each video clip was presented for 500ms. There was no interval between videos such that the first frame was presented immediately after the last frame of the prior video. If a target appeared in the video clip, it was present on each 100ms image. The distracter to target ratio was 90/10. RSVP sequences were presented in two minute blocks after which time participants were given a short break. Participants completed a total of 25 blocks.

2.3 EEG Recording and Analysis

Electrophysiological recordings were digitally sampled at 512Hz from 64 scalp electrodes arranged in a 10-10 montage using a BioSemi Active Two system (Amsterdam, Netherlands). External leads were placed on the outer canthus and below the orbital fossa of both eyes to record electrooculography (EOG). Continuous EEG data were referenced offline to the average of the left and right earlobes and digitally filtered 0.1-55Hz. To reduce muscle and ocular artifacts in the EEG signal and potential contamination with brain-based signals, we removed EOG and EMG artifacts using independent component analysis (ICA) [24].

ERP Analysis

ERP analysis was used to evaluate the trial by trial temporal variability of the neural response. Analyses for these data were previously reported [22] and are briefly described here. EEG data were processed and analyzed using EEGLAB [25] and ERPLab [26]. Continuous, artifact free data were epoched -1500 to 1500ms around target onset. Target epochs followed by a button press within 200 to1000ms and non-target epochs not followed by a response were included in the analysis. Averaging across all trials in a given condition may mask meaningful brain dynamics associated with performance; especially in perceptually difficult tasks in which the variance in ERP latency and reaction time (RT) increases [27]. Therefore, to assess the brain dynamics associated with varying levels of RT performance, target epochs were sorted into bins corresponding to an individual participant's reaction time quartile [28]. Grand averages across all subjects were then calculated for each quartile.

Single Trial Classification

The novel classification approach presented here is a modification of hierarchical discriminant component analysis (HDCA). Because of this, HDCA served as an ideal baseline measure of classification performance for this study. Details of the HDCA algorithm can be found in [7, 9–11] and it is briefly described below.

For classification purposes, EEG data were epoched -500 to 1600 ms around stimulus onset. Epoched EEG data were baseline corrected by removing the average of activity occurring between -500 and stimulus onset. Target epochs followed by a button press within 200 to 1000ms and all non-target epochs were included in the classification analysis.

Hierarchical Discriminant Components Analysis

HDCA transforms multi-channel EEG data collected over a temporal window relative to image onset into a single interest-score. Ideally, the interest score is generated so that the range of scores for each class are distinct, thereby allowing for simple discrimination of the two classes.

Generating interest scores from HDCA involves a two stage classification. In the first stage, our implementation uses a set of 15 discriminators applied to 15 non-overlapping 100 ms time windows that span 100 ms to 1600 ms after image onset. Each of the 15 discriminators is trained independently. Each discriminator combines the information contained in all 64 EEG signals collected over the course of the corresponding time window into a single value for discriminating target versus non-target. Thus, stage 1 of HDCA produces 15 interest scores that independently discriminate target from non targets. In the second stage, a separate discriminator is applied to the output of the stage 1 discriminators to create a single interest score that can efficiently discriminate between target and non-target trials.

Sliding Hierarchical Discriminant Components Analysis

Sliding HDCA (sHDCA) builds upon the standard HDCA algorithm in an attempt to extract more information from temporally scattered events. sHDCA starts by using a standard HDCA classifier trained to discriminate targets versus non targets based on 500 ms of data between 300 ms and 800 ms after stimulus onset using 50 ms time slices. Rather than simply statically applying this classifier to each epoch, in sHDCA this initial classifier slides in time such that it is applied at each sample ranging from 200 ms prior to stimulus onset to 800 ms after stimulus onset. This sliding step means that the classifier is using epoch data from 100 ms post stimulus to 1600 ms post stimulus, which matches the data used by the standard HDCA algorithm.

Because each application of the standard HDCA algorithm produces a single score, sliding the HDCA classifier in time produces a single score per application (per time point). When the sliding process is complete, we are left with a score signal that is 1000 ms in duration. From this score signal, a second HDCA classifier is trained to discriminate targets versus non-targets based on the score signal. This second level classifier uses ten 100 ms time slices. The result of this HDCA classifier is the final score assigned to the epoch which is used to decide whether the current epoch is a target or non-target.

Cross Validation

A 10-fold cross validation was used to determine the accuracy for both classification methods. Data from each subject were divided into 10 equal sized blocks of trials. Classifiers were trained on 9 of the 10 blocks, and then tested on the block left out. This process was repeated 10 times such that each of the 10 blocks of trials was used as the independent testing set once. Performance was evaluated based on the area under the ROC curve (AUC). Each participant's performance was calculated as the average AUC calculated across all 10 cross validation sets. Statistical analyses for each classification method were performed on the average AUC for each participant.

Computational Requirements

Timing measures were also employed to evaluate the computational costs of training and testing each algorithm. For this evaluation, the MATLAB functions 'tic' and 'toc' were used to measure the total time needed for classifier training and testing. The time needed for testing was divided by the total number of trials in the test set to calculate an approximation of the total time needed to apply the classifier to a single epoch as would be required in a real-time application.

3 Results

3.1 Existence of Temporal Variability

Reaction time quartiles were used as binning parameters for the ERP analysis[28]. P3 latency exhibited a large amount of temporal variability relative to the stimulus onset

(Figure 1). P3 latency data were submitted to a one-way ANOVA with the main factor of Quartile containing four levels. Analysis showed a significant main effect of Quartile, $F(3,42) = 69.37$, $p < .001$. Subsequent t-tests revealed each quartile was significantly different ($\alpha = .05$) from each other after correction for multiple comparisons using Tukey's method indicating that P3 latency increased as RT became slower. (Figure 1).

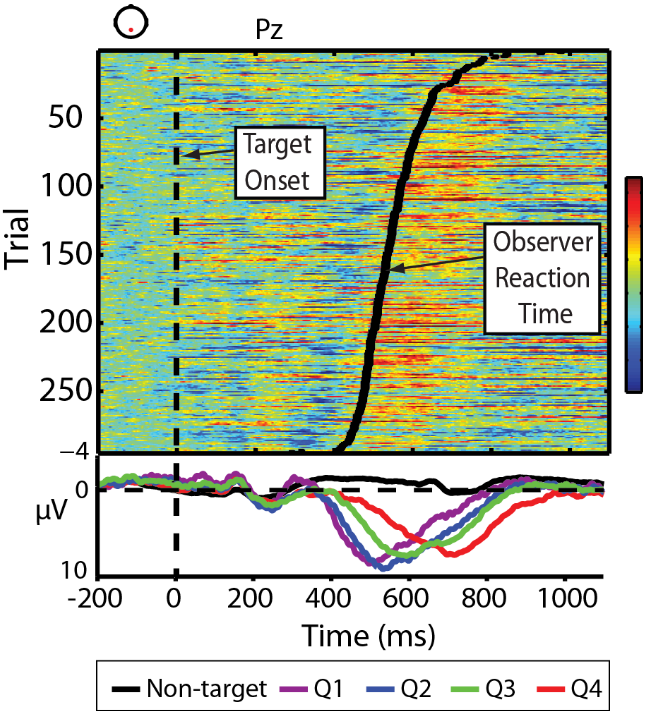

Fig. 1. Temporal variability in EEG of a single participant (S10). Upper plot shows single trial EEG response at Pz when activity is aligned to the target onset and sorted by response time. Lower plot shows average ERPs when reaction time is used as a binning parameter for ERP analysis.

3.2 Classification in the Face of Temporal Variability

Figure 1 clearly establishes the presence of temporal variability in the neural response. Figure 2 shows the accuracy of single-trial classification on these data. HDCA achieves a classification accuracy of 0.8691 ± 0.0359 (Mean AUC ± Std), while the classification accuracy of Sliding HDCA was 0.9365 ± 0.0223 (mean ± std AUC). This represents a 51.5% reduction of classification error and the overall difference is statistically different (Wilcoxon Sign Rank Test p<0.001).

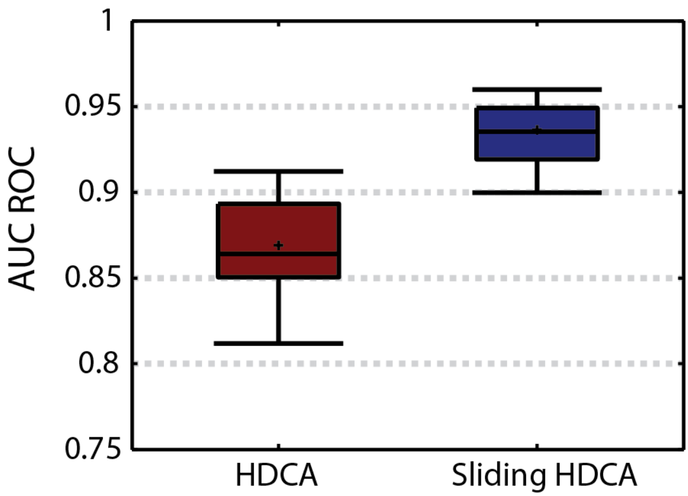

Fig. 2. Classification results across 15 subjects. Horizontal lines in each box represent the median and the dot represents the mean. The maroon box shows the classification accuracy when using the standard HDCA algorithm. The blue box shows the classification accuracy when using sliding HDCA. The difference is significant using the Wilcoxon Sign Rank Test ($p < 0.001$).

3.3 Execution Time

Sliding HDCA represents a potential improvement upon the standard HDCA classification scheme but comes at the cost of increased computing time. Training a standard HDCA classifier on the data set described here takes approximately 10 to 15 seconds. Training a sliding HDCA classifier on the same data set using the parameters described above takes 354 ± 33 seconds – a 20 to 35 fold increase in training time. Applying a standard HDCA classifier to this data set typically takes less than a millisecond per epoch, while applying sliding HDCA takes 383 ± 4 ms. While these relative time comparisons are important, in most RSVP applications, requiring approximately 6 minutes to train a classifier and 383 ms to apply the classifier is perfectly reasonable.

4 Discussion

The current study employed a dynamic RSVP task using short-duration videos. ERP analyses showed a high degree of temporal variability in the neural response. This study developed a novel classification scheme that overcame the temporal variability in the data without needing to use information from the behavioral response.

Sliding HDCA classification is a novel classification method described here that reduced classification error by over 50% over a standard HDCA classifier using the same amount of data. The increased accuracy of sHDCA classification comes at the expense of computation time. The increase in computation time is significant; however for most applications the increased accuracy seen with sHDCA will far outweigh the increase in computation time.

This study demonstrates that algorithms that account for temporal variability can dramatically improve classification accuracy. The novel method described here is one such method. This method enables further development of applications that either replace or augment behavioral responses for tasks where variable reaction times are expected.

5 Conclusion

The Sliding HDCA method described here provides a means to overcome the temporal variability in the neural response that is likely to occur in more complex environments. By transforming the raw EEG signal into a score signal, the sliding step of sHDCA produces a new signal that emphasizes the discriminating features of the EEG input and consequently improves single trial classification. The efficacy of this approach was demonstrated in an RSVP target detection task; however this approach may also prove to be useful for other types of BCI technologies in which temporal variability causes a drop in performance.

References

1. Kammer, T., Lehr, L., Kirschfeld, K.: Cortical visual processing is temporally dispersed by luminance in human subjects. Neuroscience Letters 263, 133–136 (1999)
2. Folstein, J.R., Van Petten, C.: After the P3: Late executive processes in stimulus categorization. Psychophysiology 48, 825–841 (2011)
3. Craig, A., Tran, Y., Wijesuriya, N., Nguyen, H.: Regional brain wave activity changes associated with fatigue. Psychophysiology 49, 574–582 (2012)
4. Lal, S.K.L., Craig, A.: Driver fatigue: Electroencephalography and psychological assessment. Psychophysiology 39, 313–321 (2002)
5. Magliero, A., Bashore, T.R., Coles, M.G.H., Donchin, E.: On the Dependence of P300 Latency on Stimulus Evaluation Processes. Psychophysiology 21, 171–186 (1984)
6. Dien, J., Spencer, K.M., Donchin, E.: Parsing the late positive complex: Mental chronometry and the ERP components that inhabit the neighborhood of the P300. Psychophysiology 41, 665–678 (2004)
7. Gerson, A.D., Parra, L.C., Sajda, P.: Cortically coupled computer vision for rapid image search. IEEE Transactions on Neural Systems and Rehabilitation Engineering 14, 174–179 (2006)
8. Tomioka, R., Aihara, K., Müller, K.R.: Logistic regression for single trial EEG classification. Advances in Neural Information Processing Systems 19, 1377–1384 (2007)
9. Parra, L.C., Christoforou, C., Gerson, A.D., Dyrholm, M., Luo, A., Wagner, M., Philiastides, M.G., Sajda, P.: Spatiotemporal Linear Decoding of Brain State. IEEE Signal Processing Magazine 25, 107–115 (2008)
10. Sajda, P., Pohlmeyer, E., Wang, J., Parra, L.C., Christoforou, C., Dmochowski, J., Hanna, B., Bahlmann, C., Singh, M.K., Chang, S.-F.: In a Blink of an Eye and a Switch of a Transistor: Cortically Coupled Computer Vision. Proceedings of the IEEE 98, 462–478 (2010)
11. Pohlmeyer, E.A., Wang, J., Jangraw, D.C., Lou, B., Chang, S.-F., Sajda, P.: Closing the loop in cortically-coupled computer vision: a brain–computer interface for searching image databases. Journal of Neural Engineering 8, 036025 (2011)

12. Ramoser, H., Muller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Transactions on Rehabilitation Engineering 8, 441–446 (2000)
13. Lemm, S., Blankertz, B., Curio, G., Muller, K.-R.: Spatio-spectral filters for improving the classification of single trial EEG. IEEE Transactions on Biomedical Engineering 52, 1541–1548 (2005)
14. Dornhege, G., Blankertz, B., Krauledat, M., Losch, F., Curio, G., Muller, K.-R.: Combined Optimization of Spatial and Temporal Filters for Improving Brain-Computer Interfacing. IEEE Transactions on Biomedical Engineering 53, 2274–2281 (2006)
15. Tomioka, R., Dornhege, G., Nolte, G., Blankertz, B., Aihara, K., Müller, K.R.: Spectrally weighted common spatial pattern algorithm for single trial EEG classification. Dept. Math. Eng., Univ. Tokyo, Tokyo, Japan, Tech. Rep. 40 (2006)
16. Wu, W., Gao, X., Hong, B., Gao, S.: Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL). IEEE Trans. Biomed. Eng. 55, 1733–1743 (2008)
17. Farquhar, J.: A linear feature space for simultaneous learning of spatio-spectral filters in BCI. Neural Networks 22, 1278–1285 (2009)
18. Cecotti, H., Rivet, B., Congedo, M., Jutten, C., Bertrand, O., Maby, E., Mattout, J.: A robust sensor-selection method for P300 brain–computer interfaces. Journal of Neural Engineering 8, 016001 (2011)
19. Touryan, J., Gibson, L., Horne, J.H., Weber, P.: Real-Time Measurement of Face Recognition in Rapid Serial Visual Presentation. Front Psychol. 2 (2011)
20. Yu, K., Shen, K., Shao, S., Ng, W.C., Kwok, K., Li, X.: Common Spatio-Temporal Pattern for Single-Trial Detection of Event-Related Potential in Rapid Serial Visual Presentation Triage. IEEE Transactions on Biomedical Engineering 58, 2513–2520 (2011)
21. Yu, K., Shen, K., Shao, S., Ng, W.C., Li, X.: Bilinear common spatial pattern for single-trial ERP-based rapid serial visual presentation triage. Journal of Neural Engineering 9, 046013 (2012)
22. Ries, A.J., Larkin, G.B.: Stimulus and response-locked P3 activity in a dynamic RSVP task. ARL-TR-6314 (2012)
23. Touryan, J., Gibson, L., Horne, J.H., Weber, P.: Real-time classification of neural signals corresponding to the detection of targets in video imagery. Presented at the International Conference on Applied Human Factors and Ergonomics, Miami, Fl (2010)
24. Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M.J., Iragui, V., Sejnowski, T.J.: Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37, 163–178 (2000)
25. Delorme, A., Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134, 9–21 (2004)
26. Luck, S.J., Lopez-Calderon, J.: ERPLAB Toolbox (2010)
27. Luck, S.J.: An Introduction to the Event-Related Potential Technique. A Bradford Book (2005)
28. Poli, R., Cinel, C., Citi, L., Sepulveda, F.: Reaction-time binning: A simple method for increasing the resolving power of ERP averages. Psychophysiology 47, 467–485 (2010)