A Comparison of Classification for Driver Mental Workload Using ERP and Band Power Parameters

Shengguang Lei, Sebastian Welke, and Matthias Roetting

Keywords: Mental Workload; EEG; ERP; Band Power; Driver State Monitoring

Abstract

This article outlines a study on driver mental workload classification based on two categories of information in electroencephalography (EEG) data: event-related potentials (ERPs) and frequency band powers (BPs). For data acquisition, a simulated driving task, the Lane Change Task (LCT), was combined with a secondary auditory task, the Paced Auditory Addition Serial Task (PASAT). Workload levels were manipulated by changing the speed settings and the paces of the PASAT. The ERPs and BPs (including the five frequency bands of delta, theta, alpha, beta, and gamma) were extracted from the recorded EEG data as the input for a powerful classifier, Adaptive Boosting (Adaboost). The ERP based classification showed a mean accuracy of 63% for the speed induced workload in the single task condition and a mean accuracy of 66% for the PASAT induced workload in the dual task condition. Conversely, the BP based classification showed a mean accuracy of 85% for both speed induced workload and PASAT induced workload classification. This indicates that BPs might be a better choice for the evaluation of the driver mental workload. With the new role of supervisory control in advanced humans-machine systems, there is an increasing requirement on the assessment of human mental workload and performance. Introducing electroencephalography (EEG) technology into the man–machine interface provides an effective methodology to evaluate human functional states. In the last decades, characteristic changes in the event-related potentials (ERPs) and the EEG frequency band powers (BPs) that reflect levels of workload have been identified. ERPs are the stereotyped electrophysiological responses to an internal or external stimulus, while BPs are typically extracted by dividing EEG data into frequency bands of delta, theta, alpha, beta and gamma. Raabe et al. (2005) identified a change in the amplitude of P300, a positive peak around 300ms after the stimulus onset, caused by task difficulty. Additionally, several prominent band power features in EEG data have been reported to be sensitive to the variations in mental effort. For instance, a decrease of parietal-occipital alpha rhythm and an increase of frontal theta rhythm have been observed as the mental effort of a task increases (Gevins et al., 1998). Based on these features, a number of researchers have attempted to classify mental workload using different classifiers. Ling et al. (2001) revealed an accuracy of 80% in a study to classify the simulated pilot workload levels using a neural network based on BPs. A driver's mental overload is considered to be one of the most important contributors to traffic accidents. This article presents the first results of a comparison of the contributions of two categories of candidates (ERPs and BPs) for driver mental workload evaluation. ERPs and BPs were extracted from the recorded EEG data in a simulated driving task, the Lane Change Task (Mattes, 2003) and a combination task of the LCT and the PASAT (Gronwall, 1977). A very powerful classifier, Adaboost (Freund and Shapire, 1997), was used in the present study. The comparison of these two kinds of data showed that the Band Powers could provide higher accuracy for driver mental workload classification.

Data Acquisition

In the experiment, participants were asked to perform the Lane Change Task, in which the driver is to perform a lane change amongst three lanes when prompted to by a road sign. In the single task condition, one of three speed settings (60km/h, 80km/h, and 100km/h) was used to produce workload levels. The higher speed was concluded to evoke higher mental workload (Lei, et all.,2009). For the dual task condition, participants were asked to perform a combination of the LCT at a speed of 80km/h and a two paced secondary task PASAT with the digits presented every 3 and 5 seconds (referred to as P3 and P5). The shorter interval between the digits induces a higher mental workload (for details of the experiment, see Lei, et al., 2009a). While performing the LCT the brain activity was recorded with 32 Ag/AgCl impedance-optimized electrodes (ActiCap, Brain Products). Electromyogram (EMG) was recorded from both forearms and horizontal and vertical eye-movement was recorded using the Electrooculogram (EOG). EMG and EOG signals were used for the artefacts correction in the data analysis. This way, the experiment yielded five condition datasets for each participant (three speeds, 60km/h, 80 km/h, 100 km/h, and two combined tasks,80+P5, 80+P3). These five datasets were divided into two groups, the single task (60, 80, 100) and dual task (80, 80+P5, 80+P3).

Feature extraction

EEG data analysis was performed using EEGLAB 6.03 (Delorme & Makeig, 2004), a freely available open source toolbox running under Matlab 7.3.0. For pre-processing, the data was down-sampled to 250 Hz to save computation time, and was then digitally filtered using band pass filter (1,75 Hz) to minimize drifts and a notch filter (49,51Hz) to suppress the line noise. Then the EEG data was common average re-referenced (CAR), a method which references the data to the average across all electrodes to avoid the influence of an arbitrary local reference. For artefacts correction, the EEG recordings were visually inspected for detecting the muscle artefacts and outliers. The electrodes with abnormal potential behaviour were excluded. Another powerful method, Independent Components Analysis (ICA) was used to correct the ocular artefacts (Delorme & Makeig, 2004).

Event Related Potentials

A small laplacian filter was used to re-reference an electrode to the mean of its four nearest neighbouring electrodes. The edge electrodes and some abnormal electrodes were excluded and still 18-21 electrodes remained. Then, data epochs were extracted from 2000ms before stimulus onset - the command for lane change presented on the road sign - until 2000ms after stimulus onset and the time range [-2000ms, -1000ms] was removed as the baseline. This way, an average of more then 100 epochs was extracted for each condition and each subject. Finally, ERPs were averaged every 50ms from [-1000ms, 2000ms], which yielded 60 features and therefore formed a vector of more than 1000 features as the input for the classifier from the remaining electrodes for each epoch, condition and subject.

Band Powers

For band power feature extraction, the data was epoched by a 2s time window with a 50% overlapping. This way, an average of 300 epochs was obtained for each electrode, condition and subject. Wavelet transformation was used to decompose the EEG signal into different frequency bands. Daubenchies D6 wavelet function, which has been reported to be similar to the sum of the neuron action potentials (Daubechies, 1992), was used to perform a 6-level analysis, and the output coefficients were reconstructed. This way, five EEG frequency bands delta (1-4Hz), theta (4-8Hz), alpha (8-16Hz), beta (16-32Hz), and gamma (32-64Hz) were extracted. After wavelet decomposition, the EEG band power was calculated as the sum of the square of each data point and normalized by the ratio of the power of specific frequency band to the total power of these five frequency band for each epoch. This way, these five normalized bands formed a vector of more than 100 features as the input for the classifier.

Classification

For this study, a powerful classifier, Adaboost, was used. AdaBoost, i.e. Adaptive Boosting, is a machine learning algorithm originally proposed by Freund and Shapire (1997). It offers numerous advantages. It is fast, simple and easy to program. It has no parameters to tune, requires no prior knowledge and can be flexibly combined with any method. Recently, Adaboost has been introduced in the classification of EEG signals (Pei et al., 2005). In the present study, we used decision trees based a logistic regression version of Adaboost (Friedman et al., 2000; Collins et al. 2002). Decision trees make good weak learners because they provide automatic feature selection and limited modelling of the joint statistics of data. Each decision tree provides a partitioning of the data and outputs a confidence-weighted decision that is the class-condition log-likelihood ratio for the current weighted distribution. In order to guarantee the reliability of the results, a 10-fold cross-validation was used.

Results

The average results from 20 participants are shown with the error bar chart in Figure 1. For the single task condition, ERP and BP based classification demonstrated an average accuracy of 63% (SD=14.9) and 85% (SD=10.0) respectively; for the dual task condition, the averaged classification accuracies were 66% (SD=8.1) by ERP and 85% (SD=8.5) by BP. This indicates that the BP based classifier showed much higher accuracy in both the single task and dual task condition compared with ERPs (for single task, $F(1,20)=28.95$, $p<0.001$; for dual task, $F(1,20)=53.6$, $p<0.001$).

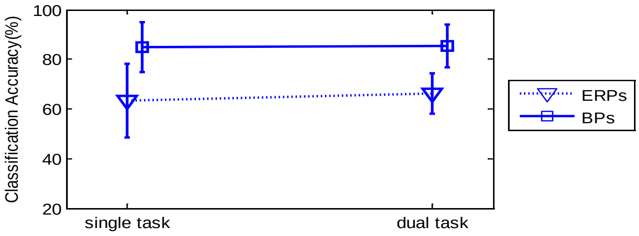

Figure 1. Mean classification accuracy of ERPs and BPs for the single and dual task condition

Discussion and Conclusion

It is well known that the ERP is the most often used feature for cognitive states evaluation. However, in the present study the ERP parameters did not demonstrate a sufficient ability to discriminate among the mental workload levels. Previous results indeed demonstrate that there are some differences in the amplitude of P3b according to the variation of the mental workload in the Lane Change Task (Lei et al., 2009a; Lei et al., 2009b). However, these differences were based on averaging more than one hundred trials. In actuality, the averaging-based ERP only involves the phase-lock information because various non-phase locked information is cancelled by the averaging. When concerning the single-trial EEG, plenty of non-phase locked information is contained. This can have great influence on classification results. Nevertheless, some methods based on spatial filtering could be used to improve classification when using single-trial ERP, e.g. the laplacian filter, Common Spatial Pattern (CSP), etc. In comparison, BPs could provide relatively more reliable results. Most of the artefacts dominate in specific frequency bands, such as ocular artefacts that are mainly included in low frequency range and muscle artefacts that are mainly found in frequencies above 15Hz. Therefore, if we could find the specific frequency bands in the EEG which are sensitive to a driver's mental workload and less contaminated by artefacts, then a reliable result could be obtained. On this point, BPs would be better suited than ERPs to deal with the artefacts, especially in situations where artefacts correction is restricted, such as real-time studies. Another limitation of using single-trial

ERP to classify the mental workload is 'real-time'. An ERP is extracted close to an event. However, mental workload is a functional state rather than an event. In the real driving environment, most of the time there are no clear events. For instance, if the driver is in a hurry or is confronted with a complex and unfamiliar road situation, there is no clear, measurable event in these situations. Additionally, it is somewhat unrealistic to use a concurrent secondary event when driving to measure the driver's mental workload. For example, some research concerning driver mental workload used an odd secondary task to detect the ERPs under different workload levels. In this context, it seems that BP is a better choice for the discrimination of driver's mental workload. To sum up, BPs may offer a better real-time feature for mental workload evaluation, and ERPs may offer a better understanding of the event-related mental activity. However, for real-time assessments of a driver's mental workload, the Band Powers might be a better choice. Future research can investigate the details of such features. Although almost 85% accuracy could be obtained, it is still quite necessary to analyze the details of the features and to ensure that no more than a few artefacts are involved in the classification. Following this, questions concerning which band is the most sensitive to driver mental workload and which area in the brain is activated by the increasing mental workload should be answered.

References

Collins, M., Schapire, R., & Singer Y. (2002). Logistic regression, Adaboost, and Bregman distances. *Machine Learning*, 48, 1-3.

Delorme, A. & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics. *Journal of Neuroscience Methods*, 134, 9-21.

Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, 28(2).

Freund Y. & Schapire R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1),119–139.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S. & Du, R. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors*, 40, 79–91.

Gronwall, D. (1977). Paced Auditory Serial Addition Task: A measure of recovery from concussion. *Perceptual and Motor Skills*, 44, 367-373.

Lei, S.G., Welke, S., & Roetting, M. (2009a). Representation of Driver's Mental Workload in EEG Data. In D. de Waard, J. Godthelp, F.L. Kooi, & K.A. Brookhuis (Eds.). *Human Factors, Security and Safety* (pp. 285-294). Maastricht, the Netherlands: Shaker Publishing.

Lei, S.G., Welke, S., & Roetting, M. (2009b). Driver mental workload assessment using EEG data in A dual task paradigm. In *Proceedings of 21st International Technical Conference on the Enhanced Safety of Vehicle*, Stuttgart, Germany.

Ling, C., Goins, H., Ntuen, A., & Li, R. (2001). EEG signal analysis for human workload classification. In *Proceedings of the IEEE SoutheastCon 2001*, Clemson, SC, 123–130.

Mattes, S. (2003). The Lane Change Task as a Tool for driver Distraction Evaluation. *IHRA-ITS Workshop on Driving Simulator Scenarios*, Dearborn, Michigan.

Pei, X., Zheng C., Xu J., & Bin G. (2005). Adaboost for improving classification of left and right hand motor imagery tasks. In *Proceedings of 2005 First International Conference on Neural Interface and Control*. 10-13.

Raabe, M., Rutschmann, R.M., Schrauf, M. & Greenlee, M.W. (2005). Neural Correlates of Simulated Driving: Auditory Oddball Responses Dependent on Workload. In Schmorrow, D. D. (Eds.), *Foundations of Augmented Cognition* (pp. 1067-1076). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. Publishers.