2009

Contribution of High-Frequency EEG Features to Physiogicallybased Operator Workload Estimation

Jason W. Monnin

Justin R. Estepp

CONTRIBUTION OF HIGH-FREQUENCY EEG FEATURES TO PHYSIOGICALLY-BASED OPERATOR WORKLOAD ESTIMATION

Jason W. Monnin
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH
Justin R. Estepp
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH

Physiologically-based cognitive workload assessment can discriminate changing levels of operator functional state in complex task environments. In this paradigm, electroencephalography (EEG) is a commonly used physiological measure. Spectral power in clinical frequency bands is used to derive features to train an artificial neural network (ANN) classifier to recognize changes in cognitive workload. Recent research has suggested that power in high frequency bands may be influenced by electromyographic artifact. In a previous study, nineteen channels of EEG were recorded (from 10 participants) during a complex uninhabited air vehicle (UAV) control simulation in which task difficulty was manipulated to induce changes in cognitive demand. In offline analysis, an ANN classifier was trained using feature sets which included and excluded features from high frequency bands. Excluding the high frequency bands reduced classification accuracy, suggesting that, while potentially of electromyographic origin, these features are still important features in physiology-based cognitive workload assessment.

Current physiologically-based operator workload estimation methods have demonstrated very high classification accuracy. Physiological data used for workload estimation typically includes electroencephalogram (EEG), in addition to electrical eye and heart activity (Wilson and Russell, 2007). However, in several recent works, it has been suggested that scalp EEG is largely contaminated by electrical musical activity from facial muscle groups (Goncharova et al., 2003; Whitham et al., 2007; Yong et al, 2008). While the exact nature of this artifact, with respect to the frequency range of contamination, is somewhat disputed (Yong et al., 2008), it is widely accepted that high-frequency components of EEG above approximately 20 Hz, especially in lateral electrode sites (furthest from the midline, are the most heavily influenced by muscle activity. A further compound of this potential problem is that, while the detection of muscle artifact through electromyogram (EMG) signals can be reliable and robust (for one example of many, see van de Velde et al., 1998), the ability to remove muscle artifact from the EEG signal is not trivial, given that both EMG and EEG share similar statistical properties and overlapping frequency content (Djuwari et al, 2005). As discovered by Fatourechi et al. (2007) in a meta-analysis of brain-computer interface (BCI) literature, the presence of muscle artifact in EEG is largely ignored by researchers.

While the artifact correction problem remains difficult, it is still necessary to be able to quantify the effect of muscle contamination in EEG-based systems. The term "artifact," when taken in the strictest sense, would tend to suggest the presence of an undesirable signal, leading to poor signal-to-noise ratio (SNR) in an available dataset. The question remains, however, whether muscle contamination of EEG is truly detrimental in any given application, or whether it could perhaps be of some benefit (given the specific nature of the application in question). The purpose of this work is to understand what effect high frequency EEG components, presumably primarily influenced by muscle artifact, have on the ability to accurately assess cognitive workload in a dynamic simulation environment. Based on the analysis presented here, the removal of high-frequency EEG components negatively impacts the accuracy of being able to discriminate between varying cognitive tasks. These results suggest that, while the source of muscle artifact may very well contaminate underlying EEG components, these high-frequency features of the EEG signal are useful in cognitive workload assessment. In addition, this work tends to agree with others who have shown increased EMG activity in a variety of muscle groups during periods of high cognitive demand (Whitham et al, 2008) and stress (Lundberg et al, 1994).

Methods

Unmanned Aerial Vehicle (UAV) Simulation Task

Ten participants (7 male, 3 female) ranging from 20 to 24 years of age (mean of 22.6, standard deviation of 4.3), after providing signed, informed consent, were trained as operators in an Unmanned Aerial Vehicle (UAV) simulation task. In this task participants were asked to simultaneously monitor multiple UAVs during four unique task conditions. Since the nature of the work presented here can be sufficiently described by considering a 2-class pattern recognition problem, data from only 2 of the 4 task conditions are presented in this analysis. Both of these task conditions consisted of 4 UAVs during a target identification and weapons pairing task. In order to increase the cognitive demand on the operator between the first and second task conditions, the air speed of the UAVs was increased in the difficult task condition, as previously used by Wilson & Russell (2007) in a similar test-bed environment. To guard against possible task learning effects during data collection trials, each participant was trained to asymptotic performance in the simulation. Asymptotic performance was defined by a minimum performance threshold of 90% target placement.

After asymptotic performance was reached, the participants returned on a separate day for a series of trials with physiological data collection. While a variety of physiological measures were recorded, for the purpose of the analysis, only EEG data are reported. EEG was recorded from 19 electrodes placed on the scalp in accordance to the International 10-20 System for Electrode Placement. The data was processed via a MicroAmps (SAM Technology, Inc.; San Francisco, CA) amplifier/filter and software package with a sampling frequency of 256 Hz, bandpass filtered from 0.5 Hz to 100 Hz. Vertical and horizontal eye artifact were corrected using an embedded linear-regression (offline) technique, where vertical electrooculogram (VEOG) and horizontal electrooculogram (HEOG) were used as representative artifact signals in the regression. Post-hoc from data collection, the data from 2 of the 10 participants were excluded from data analysis due to poor data quality.

Artificial Neural Network Training and Feature Reduction

An artificial neural network (ANN) classifier (Wilson and Russell, 2007) was trained using a feature set composed of data solely from the electroencephalogram (EEG). The EEG data was post-processed into frequency band power (via a windowed FFT of 10 seconds, with an overlap of 9 seconds) in the five traditional frequency bands (delta, 1-3 Hz; theta 4-7 Hz; alpha 8-12 Hz; beta 13-31 Hz; and gamma, 31-43 Hz) for each of the 19 electrode sites, yielding a feature set of 95 features. In order to assess the effects of the high-frequency bands on classification accuracy, the ANN was trained and tested using three combinations of frequency bands included in the feature set. The first feature set contained all 95 features, the second dataset contained the lower three frequency bands (delta, theta, and alpha, for a total of 57 features), and the third dataset included only the higher two frequency bands (beta and gamma, for a total of 18 features). A total of two-thirds of the available data was used for training the ANN (of which 25% was used as a separate validation set, independent of the training set), and the remaining one-third of the data was used as the test set. A top-down feature reduction technique was implemented using ranked saliency (calculated using a partial derivative saliency technique, as described in Greene et al. (2000)). The lowest-ranked feature, according to its calculated saliency, was removed between iterations of ANN training. The smallest feature set that yielded the highest classification accuracy on the test data was considered to be the optimally-trained network.

In order to avoid over-training and over-generalization to the provided training set, each network was verified to ensure the accuracy on the validation set (independent of the training set) was above 95%. For each network where classification accuracy was reported (the optimally-trained network), feature count totals (how many times a particular feature was used in the optimally-trained network) were compiled for each 10-20 electrode location (for all features included in a particular training set) in order to see the effects of electrode location in each training set paradigm.

Results

Mean classification accuracy (across participants) for each of the feature sets is shown in Figure 1. Using all 95 features (all bands) yielded a classification accuracy of 90.96%, which is similar to other accuracies reported in physiology-based cognitive workload studies (see Wilson and Russell (2007) for a brief overview). In the feature set containing only the three lower frequency bands (delta, theta and alpha), the classification accuracy was 76.32%. Using a feature set of the highest two frequency bands (beta and gamma), classification accuracy was

87.79%. In each feature set, the mean accuracies reported are well above chance (in a two-class pattern recognition problem chance would be 50%, like a coin-flip to determine the correct class, which indicates that the EEG features used in the training sets are sensitive to changes in cognitive workload between the two tasks).

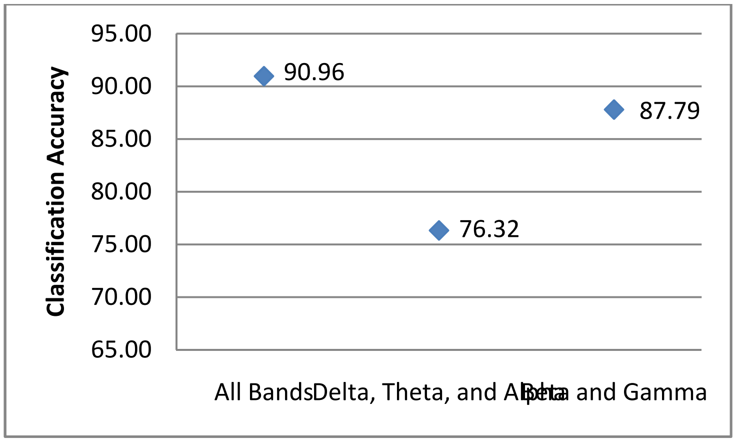

Figure 1. Mean classification accuracies (across all participants) for the three feature sets. Note the significant decrease between using all frequency bands and only using the lowest three frequency bands, as well as the increase in accuracy between using delta, theta and alpha bands verses beta and gamma.

Feature count (collapsed across 10-20 electrode sites) is shown in Figure 2. Especially in the training set where only the beta and gamma (high frequency band) features were used to train the ANN, lateral sites show up as containing the most often used features (in the optimally-trained networks) in this case, with the highest percentage originating from electrode sites: O_1, O_2 and T_4. In contrast to the all frequency band condition, the percentage of features from the medial electrode sites showed an increase in the feature set containing the delta, theta and alpha bands. In the training set containing the beta and gamma (high frequency) bands, the feature count distribution appears to be very similar to that of the all bands feature set. Lateral sites are used more often than the central electrode sites; the only centralized site with a percentage similar to the lateral sites is C_3.

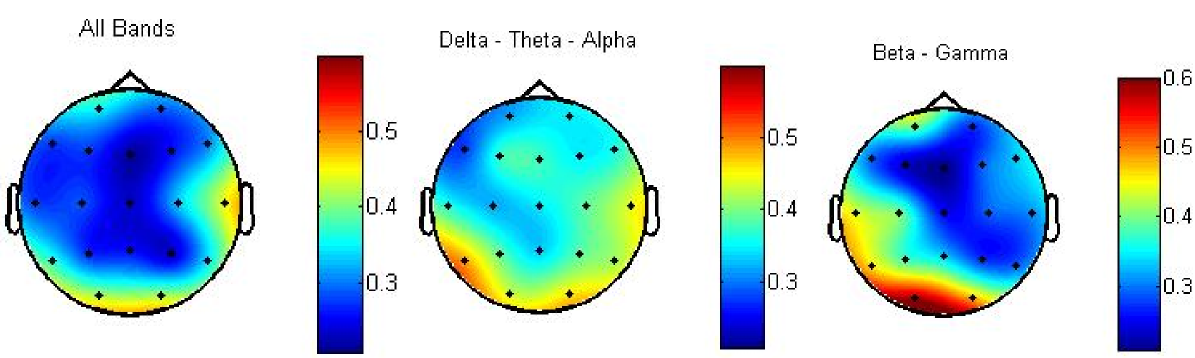

Figure 2. Feature topography using All Bands (95 features), Delta – Theta – Alpha bands (57 features), and Beta – Gamma bands (18 features), mapped to a 10-20 electrode map. Each 10-20 location shows the percentage of features from that location that were used in the optimally-trained network.

Discussion

The results obtained in the analysis agree very well with previous findings; namely, that high-frequency features in lateral electrode sites are often some of the more important features in cognitive workload classification (Wilson and Fisher, 1995). Training the ANN with the beta and gamma bands yielded higher classification accuracy over the delta, theta and alpha feature set; the difference was slightly over 11%, which demonstrates the dependency of classification accuracy on high-frequency features. However, the all bands condition (the full 95 features set) yielded the highest classification accuracy of all three feature sets. This suggests that, although higher classification is primarily driven by the high-frequency band features, the lower frequency bands still contain useful information needed to obtain the best possible classification accuracy.

The two feature sets that obtained the highest classification accuracy (the all bands condition, and the beta and gamma bands condition) also showed the highest feature selection from the lateral electrode sites on the scalp. In contrast, the delta, theta and alpha bands feature set shows that features from medial electrode sites were selected more often than either case where the high-frequency features were included in the training set. Yet, classification accuracy drops when the high-frequency features are not included. This suggests features originating from lateral sites are more likely to increase the overall classification of the ANN.

Given that the higher-frequency features are most useful in obtaining the optimally-trained network, these features are also most susceptible to muscle artifact. In addition, lateral sites are most likely to be contaminated with muscle as well, due to their location on the scalp. Since EEG above 20 Hz is largely contaminated by muscle, it suggests the all bands feature set and the beta-gamma bands feature set are contaminated with muscle as well.

Based on these results, it would appear that, specifically for the purpose of cognitive workload detection, the high-frequency features derived from EEG data are highly salient in correctly assessing levels of workload. While it is largely assumed that these features are highly driven by facial muscle artifact (especially when taken into consideration that the most likely candidate features from the high frequency bands were also from lateral electrode sites), the EEG data itself was not corrected or assessed for the presence of EMG artifact. This, in itself, suggests that EMG may be affected in a predictable manner during periods of high cognitive demand or stress, as previously mentioned in Whitham et al (2008) and Lundberg et al (1994), respectively. Therefore, this “artifact” may actually serve a useful purpose in cognitive workload classification.

References

Djuwari, D., Kumar, D.K., and Palaniswami, M. (2005). Limitations of ICA for artifact removal. *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (pp. 4685-4688). Shanghai, China: IEEE Engineering in Medicine and Biology Society.

Fatourechi, M., Bashashati, A., Ward, R.K. and Birch, G.E. (2007). EMG and EOG artifacts in brain computer interface systems: a survey. *Clinical Neurophysiology*, 118, 480-494.

Goncharova, I.I., McFarland, D.J., Vaughan, T.M., and Wolfpaw, J.R. (2003). EMG contamination of EEG: spectral and topographical characteristics. *Clinical Neurophysiology*, 114, 1580-1593.

Greene, K.A., Bauer, K.W., Wilson, G.F., Russell, C.A., Rogers, S.K. and Kabrisky, M. (2000). Selection of psychophysiological features for classifying air traffic controller workload in neural networks. *Smart Engineering System Design*, 2, 315-330.

Lundberg, U., Kadefors, R., Melin, B., Palmerud, G., Hassmen, P., Engstrom, M. and Dohns, I.E. (1994). Psychophysiological stress and EMG activity of the trapezius muscle. *International Journal of Behavioral Medicine*, 1 (4), 354-370.

van de Velde, M., van Erp, G., and Cluitmans, P.J.M. (1998). Detection of muscle artifact in the normal human awake EEG. *Electroencephalography and Clinical Neurophysiology*, 107, 149-158.

Whitham, E.M., Lewis, T., Pope, K.J., Fitzgibbon, S.P., Clark, C.R., Loveless, S., DeLosAngeles, D., Wallace, A.K., Broberg, M., and Willoughby, J.O. (2008). Thinking activates EMG in scalp electral recordings. *Clinical Neurophysiology*, 119, 116-1175.

Whitham, E.M., Pope, K.J., Fitzgibbon, S.P., Lewis, T., Clark, C.R., Loveless, S., et al. (2007). Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clinical Neurophysiology, 118*, 1877-1888.

Wilson, G.F., Fisher, F. (1995). Cognitive task classification based upon topographic EEG data. *Biological Psychology*, 40, 239-250.

Wilson, G.F., Russell, C.A. (2007). Performance Enhancement in an Uninhabited Air Vehicle Task Using Pyschophysiologically Determined Adaptive Aiding. *Human Factors*, 49, 1005-1018.

Yong, X., Ward, R.K. and Birch, G.E. (2008). Facial EMG contamination of EEG signals: characteristics and effects of spatial filtering. *Proceedings of the IEEE ISCCSP 2008* (pp. 729-734). Malta: IEEE ISCCSP.