Classification of EEG signals to identify variations in attention during motor task execution

Susan Aliakbaryhosseinabadi\textsuperscript{a}, Ernest Nlandu Kamavuako\textsuperscript{a}, Ning Jiang\textsuperscript{b}, Dario Farina\textsuperscript{c} and Natalie Mrachacz-Kersting\textsuperscript{a}

\textsuperscript{a} Center for Sensory-Motor Interaction (SMI), Department of Health Science and Technology, Aalborg University, Aalborg, Denmark

\textsuperscript{b} Department of System Design Engineering, Faculty of Engineering, University of Waterloo, Canada

\textsuperscript{c} Department of Bioengineering, Imperial College London, SW7 2AZ London, UK

*Corresponding author:
Natalie Mrachacz-Kersting
Center for Sensory-Motor Interaction (SMI)
Department of Health Science and Technology
Aalborg University
Fredrik Bajers Vej 7 D3
9220 Aalborg Ø
Denmark
Phone: 0045 9940 7571
Email: nm@hst.aau.dk
Abstract

Background: Brain-computer interface (BCI) systems in neuro-rehabilitation use brain signals to control external devices. User status such as attention affects BCI performance; thus detecting the user’s attention drift due to internal or external factors is essential for high detection accuracy.

New method: An auditory oddball task was applied to divert the users’ attention during a simple ankle dorsiflexion movement. Electroencephalogram signals were recorded from eighteen channels. Temporal and time-frequency features were projected to a lower dimension space and used to analyze the effect of two attention levels on motor tasks in each participant. Then, a global feature distribution was constructed with the projected time-frequency features of all participants from all channels and applied for attention classification during motor movement execution.

Results: Time-frequency features led to significantly better classification results with respect to the temporal features, particularly for electrodes located over the motor cortex. Motor cortex channels had a higher accuracy in comparison to other channels in the global discrimination of attention level.

Comparing with existing methods: Previous methods have used the attention to a task to drive external devices, such as the P300 speller. However, here we focus for the first time on the effect of attention drift while performing a motor task.

Conclusions: It is possible to explore user’s attention variation when performing motor tasks in synchronous BCI systems with time-frequency features. This is the first step towards an adaptive real-time BCI with an integrated function to reveal attention shifts from the motor task.

Keywords:
Attention; Attention influence; Motor movement; Global feature space; Brain-computer interface; Movement-related cortical potential

Highlights:
1. In real-world settings BCI users experience changes in attention to the main task.
2. BCI performance is significantly reduced with shifts in the user’s attention.
3. Attention to a task can be classified from EEG time and time-frequency features.
4. EEG channels located over the motor cortex provided the highest classification accuracy.
5. A General Gaussian distribution of time-frequency features improved BCI performance.
1. Introduction

Brain computer interface (BCI) systems in neuro-rehabilitation aim to help disabled people by translating brain signals into some commands to control external devices (Hallett, 1994; Terada et al., 1995). The performance of these systems is highly dependent on physiological states of the users such as fatigue (Murata et al., 2005), attention (Mangun and Buck, 1998) and emotion (Iacoviello et al., 2015). Fatigue increment, attention decrement and emotional variations may decrease BCI performance during detection of movement intention (Albares et al., 2011; Käthner et al., 2014). These parameters deteriorate the timing of neurofeedback that is a vital criterion for inducing plasticity (Argente dos Santos et al., 2012; Stefan et al., 2004). To design a robust and reliable online BCI for applications outside of the clinical environment, it is desirable to quantify the influence of these factors. To approach this aim, BCIs apply preprocessing techniques on brain signals, extract desired features and finally send a command for external device control by output of a classifier (Wolpaw et al., 2002).

Among these different parameters, we have focused on attention and we have shown in previous studies (Aliakbaryhosseinabadi et al., 2015a; Mrachacz-Kersting et al., 2015) that the attention level influences the features of EEG signals. Attention is the ability of individuals to select relevant/interesting stimuli while ignoring the other stimuli in the surrounding environment (Diez et al., 2015). Recently with an increasing interest for online BCIs, some studies have implemented techniques that have sought to identify the influence of cognitive states, such as attention, on signal properties commonly used in BCI (George and Lécuyer, 2010; Zander and Kothe, 2011). However, the effect of attention distraction during movement has not been widely explored (Antelis et al., 2012; Melinscak et al., 2016). In these studies spectral components of EEG signals were used for detection of attention focus.

Variation in attention can modulate brain signals in both the time and frequency domain (da Silva-Sauer et al., 2016; Horschig et al., 2015; O’Sullivan et al., 2015). The influence of cognitive demand on BCI performance was studied with the use of different features and classifiers in previous studies, with the purpose of translating BCI use into the natural environment (An et al., 2014; Parid et al., 2015; Schudlo and Chau, 2015). For BCI applications in rehabilitation, spectral, temporal and time-frequency features have been used to determine the alteration in the user’s state from single-trial electroencephalogram (EEG) signals (Liu et al., 2014; Lopez et al., 2009; Tonin et al., 2012; Tonin et al., 2013; Xu et al., 2014). One of the signal modalities to extract temporal features is the movement-related cortical potential (MRCP) which is a low-frequency slow cortical potential. This has been successfully implemented for movement detection and classification (Hallett, 1994; Niazi et al., 2013; Do Nascimento and Farina, 2008). The first initial negative part of this type of the control signal provides a source of information about movement preparation and user status (Aliakbaryhosseinabadi et al., 2015b; Roy
et al., 2013). In addition to temporal features obtained from the MRCP, a combination of spectral and temporal feature vectors lead to an improved classification of movement type in multi-class BCI systems (Dornhege et al., 2004; Nicolas-Alonso et al., 2015). Event-related spectral perturbation (ERSP) is one type of time-frequency feature that can be extracted for BCI control. ERSP represents the effect of a stimulus (or event) on the EEG power spectrum. A decrease or an increase in the EEG power may indicate attention variations (Akimoto et al., 2014; Jensen et al., 2007).

This study aimed at comparing temporal and time-frequency features for distinguishing the effects of imposed changes in attention on motor tasks. Time-frequency features represented a relation of spectral and temporal features as they contained spectral information in different time domains. For this purpose, time and time-frequency features were extracted from individual participants in a subject-by-subject optimized way as well as from the entire database as global features from eighteen channels located over three regions of the brain. After feature projection using the principal component analysis (PCA) method, a linear discriminant analysis (LDA) classifier was applied to discriminate normal and diverted attention status during performance of a motor task. Additionally, we intended to find a global model for attention diversion effects on the main the motor task according to the appropriate features and channels identified in the first step. For this purpose, the data from all participants were aggregated to determine a global model for feature distribution. Then these models were validated by applying them to single participant’s data.

2. Method

2.1. Experimental procedures

2.1.1 Participants

The experiment was conducted on twelve healthy participants (6 males, 6 females; mean age 24.25 ± 3.5 years). All volunteers had no history of hearing abnormality and neurological disease. The procedure was approved by the local ethical committee for the region Northern Jutland (N-20130039).

2.1.2 Experiment setup

Monopolar EEG signals were recorded from eighteen channels using an active EEG electrode system (g. GAMMAcap ¿, Austria) and g.USB amplifier (gTec, GmbH, Austria) from AF3, AFz, AF4, F3, F1, Fz, F2, F4, FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4 based on the standard international 10-20 system. The reference and ground electrode were placed on FP1 and right ear lobe respectively. Bipolar surface electromyography
(EMG) signals were recorded from the tibialis anterior (TA) muscle. All signals were sampled at 256 Hz with 16 bits accuracy.

2.1.3 Experimental protocol

Participants were seated on a comfortable chair approximately one meter away from a digital screen while their legs were placed on a step with the knee joint at 90°. The experiment contained a visual paradigm, displayed on a screen, and an auditory paradigm, which was played via a conventional headphone.

Each participant was asked to complete two tasks with different attention demands. The normal attention demand was called ‘control’ and the diversion attention level called ‘complex secondary task (CST)’. The details are outlined as below:

1. Control: In this level, the participants were asked to perform a real ankle dorsiflexion with the dominant foot timed to a visual paradigm which contained five phases of focus (2-3 sec), preparation (2 sec), task execution (0.2 sec), hold phase (2 sec) and rest time (3-5 sec). They performed 90 trials of dorsiflexion divided into three sets, each with 30 trials. Movement sets were separated by a 4-5 minutes rest period which participants were allowed to move. At the same time of movement execution, they heard auditory sounds via conventional headphones but they were asked to focus on the movement not on the sounds.

2. Complex secondary task (CST): Participants were asked to do a dorsiflexion as described above concurrently with an auditory oddball task. The oddball paradigm contained three tones called standard (500 Hz) with a probability of 60%, target (1200 Hz) with a probability of 20 % and a deviate (1900 Hz) with a probability of 20%. The sounds were played with a 75 dB sound pressure level, a five ms rise/fall time and a randomized inter-stimulus interval of 2-3 sec. As for the control level, 90 movement executions were divided into three sets with 30 trials each. Participants were asked to count the number of a special sequence of tones such as counting the number of target tone played after the standard tone while simultaneously performing ankle dorsiflexion. The type of sequences for counting was different among sets to avoid habituation.

2.2 Data analysis

2.2.1 Signal processing

Matlab software (R2014b, Mathworks®) was used to filter continuous EEG signals using a 2nd order band-pass Butterworth filter from 0.05-10 Hz to extract temporal features. EEGLAB (Delorme and Makeig, 2004) , an open source toolbox (Swartz Center for Computational Neuroscience, La Jolla, CA;
http://www.sccn.ucsd.edu/eeglab), was used for extracting time-frequency features where signals were high-pass filtered with a cut-off frequency of 0.5 Hz. Movement trials were extracted in the time window of \([-2 \, 2]\) s with respect to movement onset obtained from EMG analysis.

2.2.2 Feature extraction

The classification steps are illustrated in figure 1a. After EEG recording and signal preprocessing, two groups of features, time-domain and time-frequency domain features, were obtained from EEG trials and applied by the classifier. Figure 1b illustrates MRCP signals from different channels in two attention levels. The MRCP differed between these two attention levels, allowing the possibility to extract features from these signals for classification.
Figure 1. (A) Diagram of classification procedure. Two groups of features were compared for single-subject classification when time-domain features had better classification performance. These features were applied for global classification. (B) A sample of MRCP signals between control and complex secondary task level. It is interpreted that right lobe channels represented more differences.

2.2.2.1 Temporal features

Nine temporal features were extracted from single trials of EEG signals. The amplitude and time of peak negativity and slopes in the time domains of [-2 -1] s, [-2 0] s and [-1 0] s where 0 is the movement onset, were considered as five initial features. Variability defined as the standard deviation of each trial was also extracted in the same time ranges as the slopes. In addition, the post-movement slope obtained in the range of [0 1] s regarding to movement onset was also extracted.

2.2.2.2 Time-frequency features

Fifteen time-frequency features were extracted from different time domains of various frequency bands. A gaussian-windowed sinusoidal moving Morlet wavelet with a linear increment in the number of cycles with frequency, from a minimum of one cycle for the lowest frequency (0.5 Hz) to 20 cycles for the highest frequency (80 Hz) were applied to each single trial to obtain ERSP (event-related spectral perturbation). ERSP represented the power of the coefficients within each window. Then, ERSP values in the time ranges of [-1 -0.5], [-0.5 0] and [0 0.5] sec with regard to movement onset in the Delta (0-4 Hz), Theta (4-8 Hz), Alpha (8-15 Hz), Beta (15-30 Hz) and Gamma (30-60 Hz) bands were used as time-frequency features.

2.3 Single participant performance

2.3.1 Classification procedure
Each group of features was considered in a ten-fold test procedure to design a linear discriminant analysis (LDA) classifier [see details in (Kamavuko et al., 2015)]. In this method nine folds were used for the validation step to obtain the best LDA classifier and one remaining fold was applied to test the classifier. Following ten permutations, the results were averaged. The performance of the classification was quantified by the classification accuracy obtained from the average of the true positive rate (TPR) defined as the number of true classified points divided by the number of positive events (normal attention) and the false positive rate (FPR) defined as the portion of negative points (diverted attention) identified as positive. Dimensionality reduction using PCA was applied to the feature space prior to classification with five temporal and nine time-frequency features selected for classification.

2.3.2 Statistical analysis

The Mann-Whitney U test was used to quantify the ability of the classifier for detection of attention changes. The accuracy of the classifier was considered as the response factor while feature type with two levels (temporal and time-frequency) was the independent variable. The Kruskal-Wallis test was used to determine the effect of channel locations in representing attention alternation, with the response factor representing accuracy while the groups of channels specified within three hemispheres (right lobe, middle channels and left lobe channels) and four brain lobes (anterio-frontal, frontal, central and fronto-central lobe) were considered as independent factor. The results were considered significant when p<0.05.

2.4. Global attention threshold

We aimed to find a global criterion for attention diversion from the motor movement. According to the results from the single participant classification, time-frequency features were superior to time domain features in identifying attention level during movement execution. We thus, focused on time-frequency features from the EEG channels for the next step.

2.4.1 Features

Time-frequency features in this part were the same as for single-participant classification.

2.4.2 Global feature distribution

Since we wanted to establish a global marker for quantification of attention during task execution, the extracted features from all participants were combined to obtain a global matrix of fifteen time-frequency features for each channel. Feature classification was performed for each single channel to identify the best channel(s) as
indicator(s) for attention drift. These features were projected using the PCA method to a lower dimension space.

Nine significant features in the PCA space were selected according to the largest Eigen values. In the first step, normality of the features was tested according to the Shapiro-Wilk test using SPSS software. Then, the distribution parameters of the features such as mean and standard deviation values were obtained to design a multivariate Gaussian distribution function for these features. Finally, an evaluation test of the distribution was done with a Likelihood ratio method to classify the projected data of each participant to the same global feature space. The Likelihood ratio technique that is described with details in (Barkat, 2005), attempts to compare the goodness of fit of two models, in our experiment one of them is the normal attention state (H0) and the other one is the diverted attention status (H1). Equation (1) represents the log likelihood formula in multivariate Gaussian classification:

\[
LL(x|\mu, \Sigma) = \ln p(x|\mu, \Sigma) = -\frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma| - \frac{1}{2}(x - \mu)^T\Sigma^{-1}(x - \mu) \quad \text{(Eq.1)}
\]

Where \(d\) is the dimension of the feature space, \(\mu\) and \(\Sigma\) is the mean matrix and covariance matrix of features. (ln) represents natural logarithm.

3. Results

3.1. Single participant performance

3.1.1 Comparison of features

Two groups of features were compared according to the single participant classification. Accuracy was significantly higher in the time-frequency features in comparison with temporal features (U=30.7, \(p<0.001\)) (Figure 2a illustrates this difference). Across all participants, the accuracy of the time-frequency features was 71±10.9% while temporal features represented lower classification accuracy (65.3±8.7%).

In the next step, we sought to identify which channel locations were more affected with changes in attention.

3.1.1 Effect of channels

To determine the effect of channel locations on the detection of attention shifts, classification accuracy was analyzed based on three channel hemispheres (left, midline and right hemispheres) and also according to four channel lobes (Anterio-frontal, Frontal, Centro-frontal and Central). In this way, left hemisphere channels contained AF3,F3,F1,FC3,FC1,C3,C1, right hemisphere channels included AF4,F2,F4,FC2,FC4,C2,C4 and middle channels were; AFz, Fz, FCz and Cz. In addition, channels placed in various lobes were considered as: anterio-frontal channels; AF3, AFz, AF4, frontal lobe channels; F3,F1,Fz,F2,F4, fronto-central channels; FC3,FC1,FCz,FC2,FC4; and central channels; C3,C1,Cz,C2 and C4.
Accuracy obtained from both groups of features did not show significant differences based on channel hemispheres. On average, midline and right channels had higher accuracies than left hemisphere channels. In temporal features the accuracies of right, midline and left hemispheres were: 67.7±7%, 68.1±8.8% and 65.3±7.2%. For the time-frequency features right, midline and left accuracies were 72.7±10.7%, 71±11.1% and 69.3±11% respectively (Figure 2b).
A significant difference was found among lobes of channel placements only in accuracy according to time-frequency features ($H(3) = 8.4, p = 0.04$). Pairwise comparison revealed a significant difference between central and anterio-frontal lobe ($p=0.03$), central and frontal lobe ($p=0.04$) and also between centro-frontal and anterio-frontal lobes ($p=0.04$). The central and fronto-central channels had the best accuracy in comparison with the other lobes (Figure 2c).

3.2. Global classification

3.2.1 Feature distribution

Based on the outputs of the Shapiro-Wilk test all data in the control and attention level in the global condition had a normal distribution ($p<0.05$). Thus, mean values and standard deviations of each feature were computed to design a multivariate Gaussian distribution. Figure 3 represents a sample of the feature distributions in channel Cz in two attention conditions.

Both features had normal distribution while their mean value and standard deviation were different with regards to the attention state. Since the confidence intervals for each feature between the control and attention level were not overlapped, it was possible to separate two attention conditions by adjusting a threshold on the range of features.

3.2.2 Evaluation results

Discriminative features after PCA projection were inserted to the feature distribution and classified based on the log Likelihood ratio. Accuracy obtained from the evaluation of the classifier was compared according to the channel placement and only the central channel lobe had significant effect on attention classification ($F(3,212)=16.2, p<0.001$). Post-hoc analysis demonstrated that there are significant differences between the central and frontal lobe ($p<0.01$), central and anterio-frontal lobe ($p<0.001$) and also between fronto-central and anterio-frontal lobe ($p=0.01$).

4. Discussion

In the first part of this study, temporal features extracted from the MRCP were compared with time-frequency features obtained from EEG signals according to the different levels of attention to the motor task in anterio-frontal, frontal, fronto-central and central channels to investigate BCI performance under attention variations during movement execution. Results suggest that time-frequency features are more reliable than time-frequency domain features (B) Effect of channel hemisphere on classification accuracy (C) Effect of channel lobe on classification accuracy. The significant levels are represented by (*). It is concluded that by using time-frequency features, effect of channel lobe is more than channel hemisphere when attention to a task is diverted.
Figure 3. Distributions of features 1 and 2 in channel Cz for two attention levels in the (A) Control level and (B) CST level. The feature spaces are separable between the two attention states as the overlap of feature ranges is not significant. The mean value and standard deviation of these features show different ranges for the two attention levels.

features for modulation of attention level during performing of motor task. In the second part, the global distribution function was found for PCA projected features in two attention states to evaluate the criteria reflecting attention diversion. As for the result for single participant, channels located in the fronto-central and central regions have a significantly better performance than the other lobes.
4.1 Single-participant classification

The time-frequency features from the MRCP which were used for attention classification during motor task execution, were statistically different. Time-frequency features were obtained from different band power and different time blocks of EEG signals which have been used in previous studies on attention (Kim et al., 2015; Polomac et al., 2015). Better performance of tempo-spectral features is in line with previous studies that show superior movement detection performance with spectral or tempo-spectral features (Jochumsen et al., 2015; Kamavuako et al., 2015). These types of features represent a combination of time and frequency properties and thus contain more useful information for understanding the effect of attention on motor task performance. The classification performance of attention states is also comparable with previous studies using tempo-spectral features for movement classification under variations of task related parameters such as force level (70-80%) (Farina et al., 2007; Gu et al., 2009).

According to the results of single channel analysis, we found that fronto-central and central lobes are the most affected parts of the brain under attention diversion from the main motor task. The central and fronto-central lobe demonstrated the best classification performance in comparison to the frontal lobe. Although this is partially in contrast with some previous studies which showed that the activation of the frontal lobe is increased due to dual-tasking in comparison to the other parts of the brain (Contreras-Rodríguez et al., 2015; Mirelman et al., 2014). This may be due to a cognitive response to the more complex dual task conditions which enhance the level of oxygenated hemoglobin (HbO₂). In contrast, our results are in line with Wu et al. (2013) which suggest that the type of the tasks plays a key role in functional brain connection. Thus channels located over the motor cortex corresponded to a better performance compared to those located over the other lobes, presumably because this brain region provides the final output command to control motor activities (Fall and de Marco, 2008; Kleim et al., 2003).

Although there was no statistically significant difference among channels located in the right, midline and left hemisphere, midline and right channels had higher attention detection accuracy during motor task execution. This is supported by previous studies that showed an increased activation of the right hemisphere with auditory spatial analysis (Weeks et al., 2000; Weeks et al., 1999). In the current study the classification accuracy for channels located in this hemisphere was significantly better than for those located in the other hemisphere (Figure 2). Thus it is likely that the right hemisphere was influenced to a greater extend by the auditory stimuli and thus significantly affected the attention to the main task (dorsiflexion).

4.2 Global attention classification during motor task execution
The results of classification using the global time-frequency feature distribution showed that channels placed on the fronto-central and central lobes had the best performance in representing of attention level in motor task execution. Central channels have previously been shown to be the most optimal for motor movement classification (Miller, 2012; Yazawa et al., 1997). Better performance of these channels was predictable because upper and lower limb motor movements are mainly planned and controlled by connectivity of motor cortex regions (Crone et al., 1998; Volz et al., 2015). The effect of auditory selective attention can be monitored in channels of the same lobe in addition to the auditory cortex (Galbraith et al., 2003) as our results support this notion.

The accuracy of the global model obtained here suggests that it is possible to define a global criterion for investigating attention levels during motor task execution and to use this in online BCI systems for detection of plastic changes. Since plasticity induction is modulated with attention shifts (Conte et al., 2007; Stefan et al., 2004), a BCI for neuromodulation must incorporate an element where attention states are detected. In this way the user can be provided with the appropriate neuro-feedback to direct attention back to the main task. The global model for attention diversion from the motor task can be implemented to avoid training the classifier for different participants. This is important in clinical applications to reduce to a minimum the time spent in user training.

4.3 Study limitation

In this study, continuous EEG signals from 12 healthy participants were analyzed offline. It is vital to have online BCI systems since it is required for appropriate feedback when attention is drifted during BCI use. Furthermore, it is worthwhile to perform this experiment for patients such as stoke patients. The other limitation was the combination of motor movement and attention because attention was diverted while performing ankle dorsiflexion. So, it was not possible to analyze the effect of attention without considering the influence of dual tasking, although the main aim of this study was to quantify classification performance under attention diversion during task execution.

Conclusion

In the first part of this study, we aimed to explore the more reliable feature type from selected channels for attention classification during the execution of a simple motor task in BCI systems in healthy participants. Temporal and time-frequency features were extracted from channels in three different brain lobes. The results revealed that time-frequency features obtained from channels located over the motor cortex region have a significantly better performance for attention classification. In the second part of this paper, we aimed to find out
a global distribution for significant features of classification and then validate these functions on single participants. Multivariate normal distribution of features demonstrated the best performance in the channels located over the motor cortex.

Acknowledgements

This work was supported by grants from Det Obelske Famaliefond.
References


, 2015. Attentional selection in a cocktail party environment can be decoded from single-trial EEG. Cereb.
Cortex, 25, 1697-1706.

activity prediction: A survey. in: Chen, Y.P.P., Current Bioinformatics, Australia, 10, 344-359.

and connectivity of the early auditory evoked gamma band response. Brain Topogr., 28, 865-878.

Roy, R. N., Bonnet, S., Charbonnier, S., Campagne, A., 2013. Mental fatigue and working memory load
2013, 6607-6610.


visuospatial attention. J. Neural Eng., 9, 045011-.

Tonin, L., Leeb, R., Sobolewski, A., Millán, J. d. R., 2013. An online EEG BCI based on covert visuospatial

connectivity during movements of the upper and lower limbs. NeuroImage., 119, 44-53.


positron emission tomographic study of auditory localization in the congenitally blind. J. Neurosci., 20, 2664-
2672.


