A Novel Technique for Selecting EMG-Contaminated EEG Channels in Self-Paced Brain-Computer Interface Task Onset

YoungJae Song[1], Francisco Sepulveda[1]

[1] BCI-NE Group – School of Computer Science and Electronic Engineering, University of Essex, United Kingdom

E-mail: yjsong.bci@gmail.com

Abstract

Objective: Electromyography artefacts are a well-known problem in Electroencephalography studies (BCIs, brain mapping, and clinical areas). Blind source separation (BSS) techniques are commonly used to handle artefacts. However, these may remove not only EMG artefacts but also some useful EEG sources. To reduce this useful information loss, we propose a new technique for statistically selecting EEG channels that are contaminated with class-dependent EMG (henceforth called EMG-CCh). **Methods:** The EMG-CCh are selected based on the correlation between EEG and facial EMG channels. They were compared (using a Wilcoxon test) to determine whether the artefacts played a significant role in class separation. To ensure that promising results are not due to weak EMG removal, reliability tests were done. **Results:** In our data set, the comparison results between BSS artefact removal applied in two ways, to all channels and only to EMG-CCh, showed that ICA, PCA and BSS-CCA can yield significantly better ($p<0.05$) class separation with the proposed method (79% of the cases for ICA, 53% for PCA and 11% for BSS-CCA). With BCI competition data, we saw improvement in 60% of the cases for ICA and BSS-CCA. **Conclusion:** The simple method proposed in this paper showed improvement in class separation with both our data and the BCI competition data. **Significance:** There are no existing methods for removing EMG artefacts based on the correlation between EEG and EMG channels. Also, the EMG-CCh selection can be used on its own or it can be combined with pre-existing artefact handling methods. For these reasons, we believe this method can be useful for other EEG studies.

Keywords: Artefact removal; Blind source separation; Brain-Computer interface; EEG; EMG/EOG artefacts; EMG-CCh selection; CCA; ICA; PCA;

I. INTRODUCTION

Electroencephalography (EEG) is the most common recording modality in brain-computer interfaces (BCIs). Its signals can easily be recorded with non-invasive electrodes on the scalp. As a result, the brain signals have to cross many layers of tissue between the source of activity and the sensors. For this reason, EEG signals usually have low signal-to-noise ratios and usually contain electromyography (EMG) and/or electrooculography (EOG) artefacts [1]. These artefacts are normally larger than the EEG signal, so contamination is a well-known problem not only for BCIs but also for brain mapping and clinical EEG studies [2-4]. However, in an EMG artefact BCI review [5], Fatourechi et al. found that 67.6% of the investigated BCI studies did not mention whether EMG artefacts were removed and 12.1% clearly did not remove EMG artefacts. 10.9% of the studies were found to have manually removed EMG artefacts, which can only be applied in offline systems. 6.2% used automatic rejection (mostly thresholding methods) and the rest, 3.2%, used automatic EMG removal method such as PCA and ICA [5].

Previous BCI and EEG studies applied artefact handling techniques to all EEG channels. As such, e.g., in the case of artefact removal using blind source separation, a common approach in BCIs, there may be significant loss of useful EMG-free EEG information [6-9]. In order to minimise this information loss, we propose an algorithm for rejection of EEG channel that are contaminated by class-related EMG artefacts, channels which have been called EMG-CChs. Typical EMG handling techniques such as ICA (Independent Component Analysis), PCA (Principal Component Analysis) and BSS-CCA (Blind Source Separation-Canonical Correlation Analysis [6, 10]) will only be applied to EMG-CCh, not to all channels, in this study. This combined approach was then compared with existing methods without doing channel elimination, i.e., methods that apply EMG artefact removal to all EEG channels. The comparisons were done with our own onset detection data as well as with a BCI competition data set. To the best of our knowledge this kind of EMG artefact contaminated channel selection approach has not been presented in previous EEG studies.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2018.2847316, IEEE Transactions on Neural Systems and Rehabilitation Engineering

II. Methodology

A. Experimental Paradigm and Data Set Description

In this paper two different experiment data sets were tested. One set was from our covert and inhibited overt sound-production onset detection study [11, 12] and the other was from BCI competition IV data set 2a [13]. Both sets are discussed below.

i) Sound-production Related Cognitive Tasks for Onset Detection: In that study, four different mental tasks were tested for the onset detection: Inhibited-overt high tone; inhibited-overt siren-like sound; covert high tone; and covert siren-like sound. These four different sound production cognitive tasks were classified against the idle (a.k.a. non-control or null) state as an onset switch.

The study had seven subjects. Thus, the total number of runs was 28 (7 participants * 4 tasks). Each task run had 30 trials. In a single trial, there were 3-30 seconds of idle state followed by 3 seconds of an intentional command (IC, i.e., one of the four cognitive tasks) state. The idle state length was freely chosen by the user as the study was focused on self-paced activation of the BCI system. The recorded EEG data were segmented using a 0.5s window without overlapping and were then separated into idle and task state categories. These separated idle and task segments were pre-processed and applied with various feature extraction method. Then, Davies-Bouldin index (DBI [14, 15]) was calculated to test class separation. There were 64 electrodes placed based on 10-10 layout system with 2 reference channels on both left and right earlobes. In addition, 4 facial electrodes were placed to detect EOG and EMG artefacts. **EMG_ch1** was located above the corrugator muscle and it was used mainly to detect forehead EMG and eye blink artefacts. **EMG_ch2** was set above the levator labii and zygomaticus muscles to detect facial twitches and upper limbs artefacts as well as up/down EOG, conjointly with **EMG_ch1**. **EMG_ch3** and **EMG_ch4** were placed around the anterior-most edge of the temporalis muscle to detect left/right EOG and temporal EMG artefacts. A Biosemi ActiveTwo system was used with 512 sample/s. (more detailed descriptions can be found on [11, 12]).

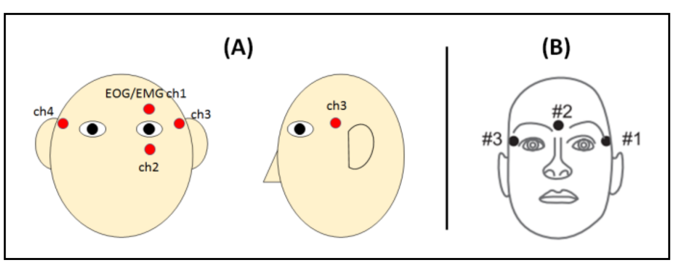

Figure 1. (A): Four facial EOG/EMG electrodes placement for our onset detection system. (B): Three facial electrode channels for BCI competition data set [13].

ii)* BCI Competition Data:** The BCI competition IV data set 2a was used (9 subjects, four different motor imagery tasks in a cue-based system) [13]. In order to simulate a self-paced scenario using these data, we only selected specific time segments. The first 0-2s segment was the preparation state (fixation cross) that corresponded to the idle state in our data. However, there was a short acoustic warning tone at 0s, which could produce event-related potentials. In addition, the length of fixation cross state is always 2 seconds, which participants can anticipate. Thus, we defined this state as ***non-specific / expectation state (NE state). However, the NE state would be expected to be somewhat similar to that preceding a self-paced task event. For this reason, we treated the NE state as corresponding to the idle state. Caution is needed when analysing results stemming from this, but the choice is still suitable for this study as the data set was chosen merely to extend our analysis to include an existing and well accepted data set. Also, this data set was chosen as it was the only BCI competition set that includes facial channels (three EMG facial channels), which were needed for a suitable comparison with results from our data. The four different motor-imagery tasks were regarded as being in the ***Task state*** class. To minimise ERP-related effects, we only used the 4-6s segment for the Task state (details in [13]).

B. Spectral Domain EMG Artefact Content

Spectral domain EMG artefacts were analysed to find out whether these artefacts affect class separation and, if so, in which frequency range. This analysis was done with our self-paced onset detection data set [11, 12]. In the analysis, the frequency bands were separated into eight different ranges. **Freq1**: 2-4Hz (Delta), **Freq2**: 4-8Hz (Theta), **Freq3**: 8-12Hz (Alpha), **Freq4**: 12-16Hz (Low Beta), **Freq5**: 16-20Hz (Beta), **Freq6**: 20-30Hz (High Beta), **Freq7**: 30-42Hz (Low Gamma) and **Freq8**: 42-100Hz (High Gamma).

Figure 2 illustrates the absolute Pearson correlation coefficient (CRC) values between 64 EEG channels and EMG channel 1 for each of the eight frequency bands. The CRC values between the filtered time series EEG and EMG data were calculated for 0.5s windows and the values shown on the figure were averaged across all subjects, all four mental tasks, and all trials, for illustration purposes. Panel (A) shows data from the idle state while panel (B) shows data from the IC task state. Also, all data was submitted to the eye blink and EOG artefact removal procedure (explained below). Thus, the data shown contains only EEG and EMG artefacts. Red and orange areas indicate EEG channels that are highly correlated with EMG channels, i.e., they had high contamination by EMG artefacts. As expected, the plots show that frontal areas have EMG contamination as seen by the high correlation with EMG_ch2 throughout all the frequency bands. On the other hand, EMG_ch3 and EMG_ch4 showed high correlation with temporal area EEG.

In [16] it was shown that EMG can contaminate all EEG bands and its contamination level differs with scalp location. We observe a similar pattern in our data. In addition, EMG contamination for the idle (panel A) and IC (panel B) states shows quite similar results for both cases. This indicates that EMG artefacts can appear in the idle state as well as in IC states. I.e., **EMG contamination per se may not be idle vs. IC class-dependent**. For this reason, we statistically checked which channel locations were affected by EMG artefacts during IC states more than during idle states. Such channels (i.e., EMG-CCh) were removed in the case of the channel rejection method. In the case of the EMG artefact removal, on the other hand, they were further processed with other EMG handling procedures. Thus, this artefact handling procedure ensures that performance results are reliable as the artefact related class-dependent effects are eliminated. The details of the procedure for EMG-CCh selection method will be discussed in section *D* below

C. EOG Artefacts Removal

EOG artefacts can be separated into two different types. One is eye blink and the other is eyeball rolling. Eye blink artefacts can easily be detected as they have relatively high signal amplitude compared to EEG. However, eyeball rolling EOG is somewhat different. Thus, in this paper, eye blink EOG artefacts were rejected with a discrete wavelet transform (DWT) denoising method while eyeball rolling artefacts were handled with the EMG & EOG contaminated channel selection method.

The window segments that contained the eye blink artefacts were automatically rejected based on the method shown in [17, 18]. A DWT (Discrete Wavelet Transform) was applied with Haar mother wavelet as it resembles eye blink artefacts. The signal was decomposed up to level 6 and was then reconstructed using only the approximation coefficients, as in [17, 18]. If the reconstructed signal's standard deviation (std) was higher than 3 times the preceding data's std (using the previous 500ms window), then this data segment (all channels) was regarded as eye blink artefact and was rejected from further analysis.

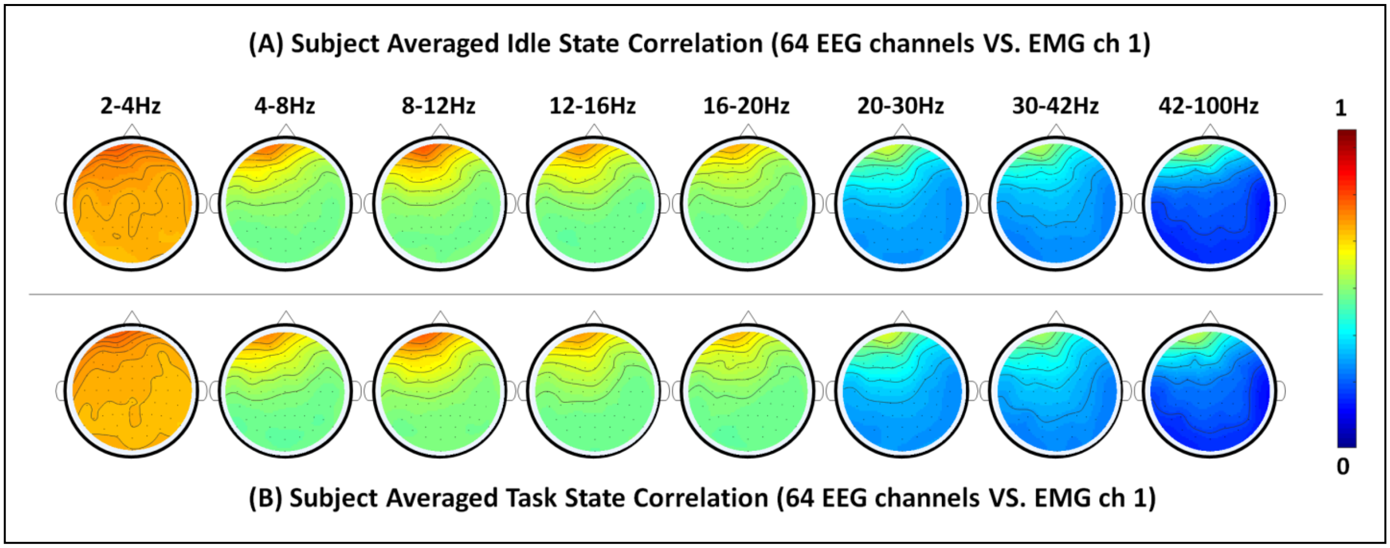

Figure 2. Topographic map of absolute Pearson correlation coefficients between 64 scalp EEG electrodes and the EMG channel 1 (EMG_ch2, 3 and 4 were not included due to space limit but it showed similar pattern with EMG_ch1). The correlation values in each channel were averaged over all subjects, four onset tasks and all trials. Panel (A) represents during idle state while (B) shows during IC task state. Freq1-Freq8 ranges from section *B* first paragraph.

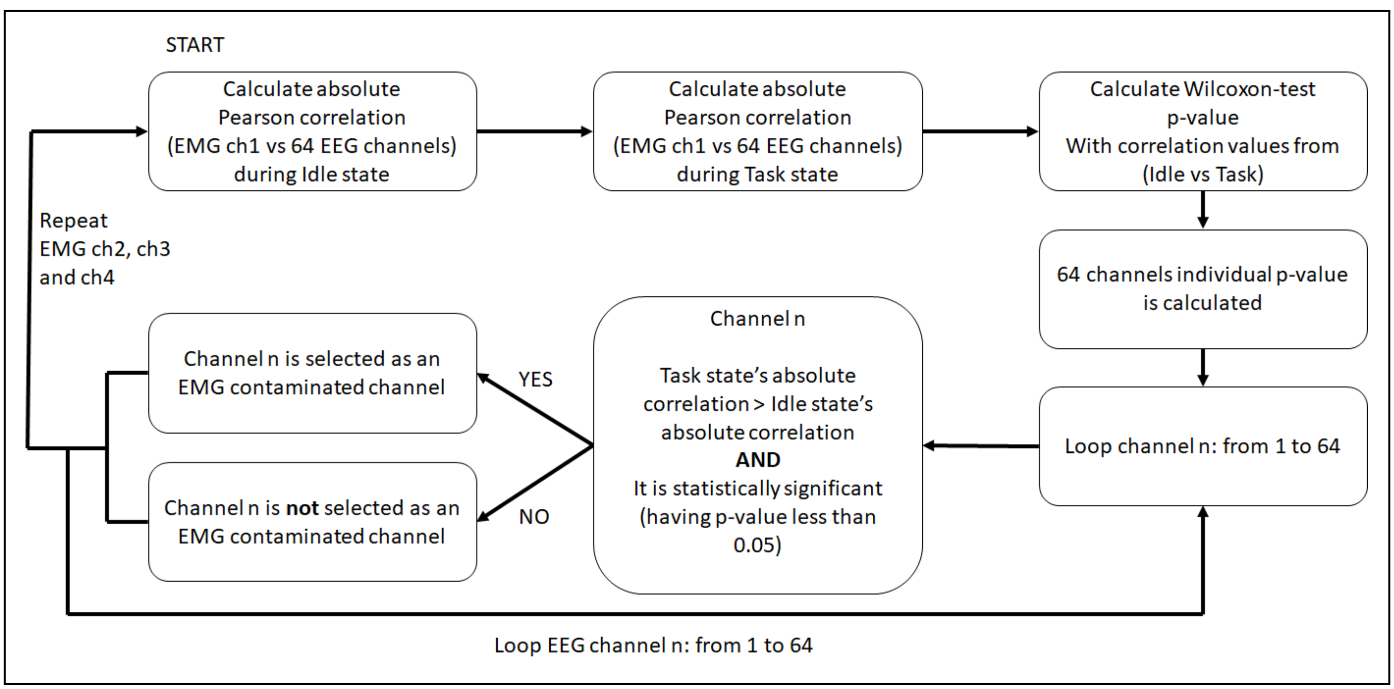

Figure 3. EMG artefacts contaminated channel selection procedure.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2018.2847316, IEEE Transactions on Neural Systems and Rehabilitation Engineering

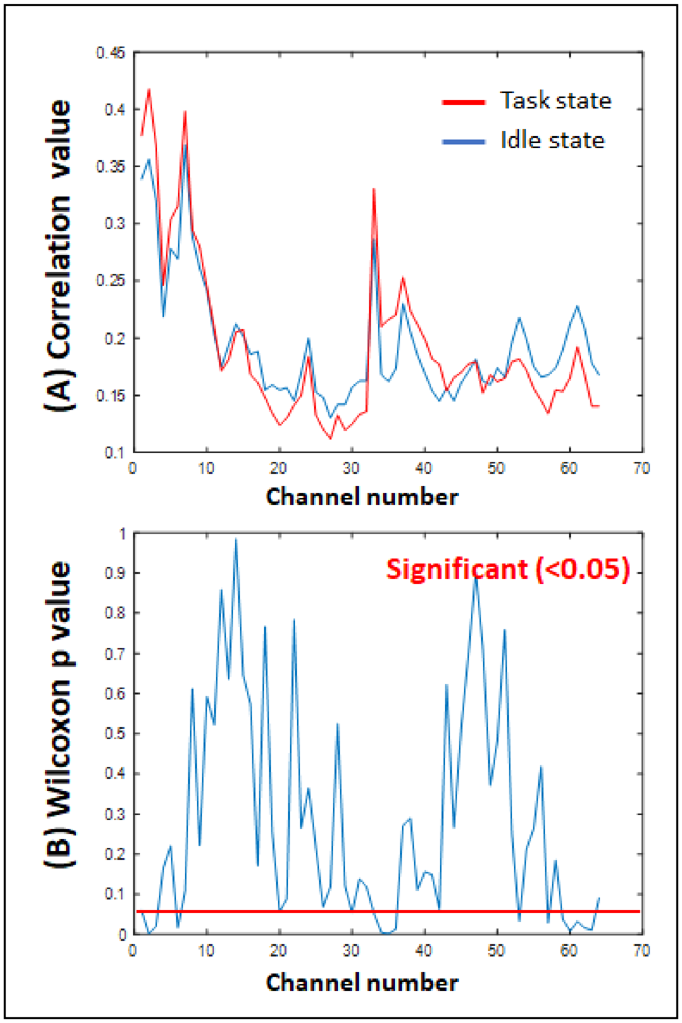

Figure 4. An example of EMG-CCh selection procedure for Participant 1's inhibited overt siren task. Figure (A) represents correlation values from 64 scalp EEG channels vs. EMG ch1. Figure (B) shows Wilcoxon test p value between idle and task state's correlation values.

D. EMG Artefact Channel Selection

The method described in this section comprises the main novelty in our study. The technique can be applied as an EMG-contaminated channel rejection method on its own, or it can be combined with other EMG handling algorithms such as ICA, PCA, BSS-CCA, etc. By applying the latter EMG handling methods only to the EMG contaminated channels, useful signal information loss can be reduced. Figure 3 shows the EMG-artefact contaminated channel selection procedure:

1) Calculate the absolute Pearson correlation (CRC) values between 64 EEG channels and EMG_ch1 for each idle and task state (0.5s time windowed filtered data). This will generate *64*NI window segments* idle state correlation values and *64*NT window segments* task state correlation values, where *NI* and *NT* are the numbers of idle and task trials, respectively.

2) Calculate the Wilcoxon-test *p*-values between the two lists (*NI* and *NT*) for each of the 64 channels.

3) If an EEG channel's EMG correlation value for the task state (average of *M* correlation values) is higher than for the idle state (average of *N* correlation values), AND the difference is statistically significant ($p<0.05$), THEN this channel was selected as an EMG contaminated channel, EMG-CCh.

4) Repeat steps 1 to 3 with EMG ch_2, EMG_ch3 and EMG_ch4.

Rationale of the procedure:

The aim of the method is to select EMG-artefact contaminated channels that affect classification (idle vs. IC) results. Participants were instructed to stay calm and relaxed in the idle states. However, when they executed a task state, unexpected EMG artefacts (e.g., facial twitches and eye movements) can contaminate EEG, especially if they are unfamiliar with BCI experiments, even though they were instructed to avoid any muscle movement. Thus, it is difficult to identify whether the result is purely EEG-based or it is EMG-contaminated. Many other blind source separation EMG removal methods have been suggested to deal with this issue. However, as existing methods are applied all EEG channels, it is possible that important EEG information is being lost in the process. For this reason, our EMG-contaminated channel selection method calculates correlations between EMG channels and EEG channels, and contaminated channels are selected for applying other EMG handling methods.

Possible limitation in our method: It is possible that a channel with lower CRC in idle states could still have too much class-dependent EMG. To test whether this may have been the case, a reliability analysis was performed (see section *H* below).

This EMG-CCh selection procedure is quite strict, unbiased and reduce valuable data loss when compared with typical thresholding methods presented in [5]. Figure 4 shows an example of the proposed EMG-CCh selection process. Figure 4A represents correlation values from 64 channels for each idle and task state (averaged from number of window segments). Channel number 35, for example, shows task state's EMG correlation is around 0.23, i.e., it is unlikely to be correlated with EMG artefacts if we cut it based on a typical thresholding method. However, if it is statistically compared with idle state by using the *p*-value from Figure 4B, it is certain that the task state's EMG correlation is significantly higher than in idle state. This indicates that channel number 35 is affected by EMG during the task activation. Thus, this channel has to be processed with further EMG handling methods to remove class-dependent EMG related performance results.

Figure 5 shows the artefact-contaminated areas from the EMG-CCh selection method for each participant and task. The red area represents the EMG-CCh. In some cases, few or no channels were selected. On the other hand, 50 channels were selected as EMG-CCh in the worst case (participant 1 C_Siren task). The EMG affected channels are task-dependent and it are not consistent between subjects. It could be due to the fact that the level of tasks activation depends on the subject and on the mental task. However, this strict EMG handling processing ensures that all the remaining analysis on EEG- based class-dependent information.

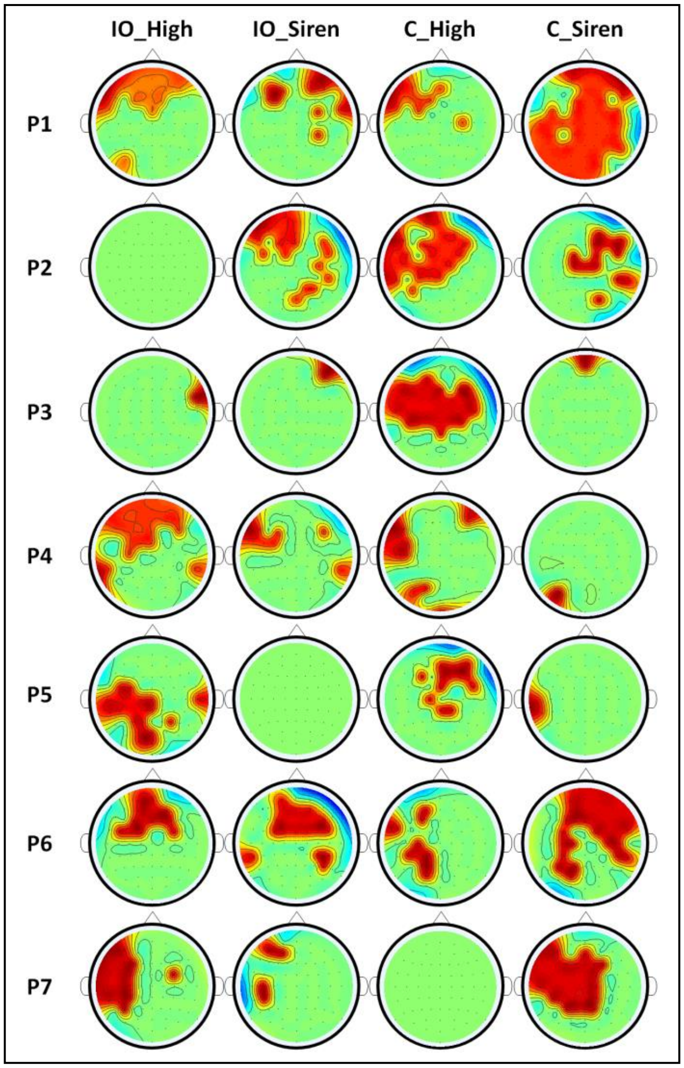

Figure 5. EMG-contaminated channels selected in the EMG-CCh selection method (red: contaminated – value '1', green: uncontaminated – value '0'). The orange and blue areas have no meaning. They are simply caused by the colour map drawing algorithm.

In BCI competition data set, Participant 2 and 3 had all the twenty-two channel selected as EMG-CCh. This means that their data were so contaminated by EMG as to preclude their use. Participants 5 and 8 had no EMG-CChs. Thus, these four subjects' data were not used for further analysis as this did not allow a comparison between typical EMG handling techniques and our method.

E. EMG Artefacts Handling

After the EMG-CCh selection process, various common EMG artefact handling methods were tested.

- ***Simple Channel Rejection*:** EMG contaminated channels were simply eliminated from further analysis.
- ***Blind Source Separation Canonical Correlation Analysis (BSS-CCA)*:** In this paper, the threshold for the autocorrelation coefficient ρ was chosen as 0.35 based on the study in [19]. Thus, EEG sources that had $\rho < 0.35$ were removed. If there is no source that has less than the threshold ρ value, the last source that has the lowest autocorrelation coefficient was removed.
- ***Independent Component Analysis (ICA)*:** In this paper, automatic artefact ICs were detected using Kurtosis and Entropy, which was suggested in [20, 21]. Both were computed for all the ICs and were normalized to a 0 mean and 1 standard deviation. The threshold was set at ±1.64 (based on [20]). If the IC exceeded the threshold, it was regarded as an artefact component and was removed.
- ***Principal Component Analysis (PCA)*:** In this study PCs that accounted for 95% of the total power were used (based on [22]) and the remaining PCs were removed.

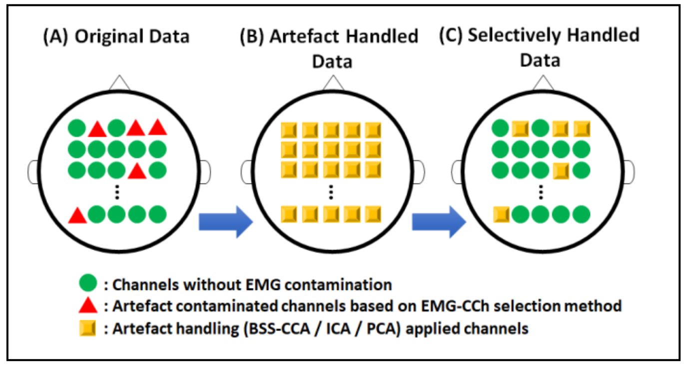

Figure 6. Examples of artefact-contaminated channel handling procedure.

These typical artefact handling techniques in BCIs were applied only to the EMG-CCh selected in the earlier stages of the analysis. Figure 7 shows an example of how to selectively apply EMG handling methods to those EMG-CCh Panel (A) shows the original data. The green circles represent normal channels that have no class separation effect from artefacts, based on the statistical test. The red triangles depict channels with EMG/EOG artefact contamination. The artefact handling methods (BSS-CCA, ICA or PCA) were applied to all the channels (64 for our onset detection data and 22 for the BCI competition data). This is represented in panel (B) with yellow squares. Panel (C) shows the channel data that were used for the final analysis. Channels without artefact contamination were kept and only the EMG-CChs were selected from (B). BSS-CCA, ICA and PCA separate N number of source components if they have N number of channels. Thus, applying these techniques only to EMG-CChs, which sometimes could be in very low numbers of channels, would not be suitable to separate artefact components with normal blind-source separation methods (although, some literature showed single channel artefact removal [23] but it is out of scope for this study). For this reason, existing artefact handling processes were applied to all channels first and then selectively taking channels if they were EMG-CCh (i.e., only EMG-CCh were replaced by the BSS cleaned channels). As a result, panel (C) has EMG-free data (i.e., statistically no artefact-related class-dependent effects), which were then used for further analysis.

F. Feature Extraction

The artefact-free data were submitted applied to four separate feature extraction methods, as follows:

- ***Autoregressive Model (AR)*****:** AR model extraction was applied to all channels and the obtained coefficients were used as features. For the coefficient estimation, Burg's method [24] was used. In [25], AR model order number 6 was suggested as optimal for imagined speech EEG signal, so order 6 was chosen.
- ***Band Power (BP)*****:** Fast Fourier Transform (FFT) was applied to the pre-processed signals with seven different frequency ranges (4-8Hz to 42-100Hz as described in *'Spectral Domain EMG Artefact Analysis' above*). Freq1 (Delta band) was removed as it highly correlates with EMG channels (as seen in Figure 2). Each bands FFT was squared and these were used as features.
- ***Common Spatial Pattern (CSP)*** [26]**:** Using our data idle and sound-production related states were used as the two separate classes. In BCI competition data set, on the other hand, four motor-imagery tasks were regarded as one class and non-specific states for the other class. After the spatial filter process, EEG source components were sorted to maximise variance for one class and to minimise it for the other class. CSP makes the first and last components represent the maximum variance difference between classes. Thus, the first three and the last three EEG source components were found. Linear regression was applied and its slope was used as a feature.
- ***Discrete Wavelet Transform (DWT)*****:** Pre-processed data were decomposed up to level 7 for sound-production related data and up to level 6 for the BCI competition data (due to different sampling rate). Then, the coefficients for the detail components, which represent pseudo frequency bands 4-8Hz, 8-16Hz, 16-32Hz, 32-64Hz and 64-128Hz (up to 100Hz as it was bandpass filtered), were obtained and the variances of the coefficients in each band were used as features. The mother wavelet 'db2', which has just 4 coefficients, was chosen because of its simplicity and because it is commonly used in BCI studies. In our previous study ([27]) we also showed that the choice of wavelet type (db2, coif2 and sym2) did not have a significant effect in onset detection in our (covert) sound-production scenario.

These features were used for Davies-Bouldin Index (DBI) calculations for evaluation purposes (details are given below). This DBI value can be different depending on which feature domain is analysed. It is for this reason that various feature extraction methods commonly used in BCIs were tested in order to increase the reliability of the evaluation. In addition, the choice of the feature extraction methods listed above covers the time, frequency, spatial and time-frequency domains.

G. Evaluation and Feature Selection

In order to evaluate how the EMG-CCh selection and handling method improves class separation compared to the common EMG handling techniques, DBI was used. DBI is a cluster separation measure that was suggested in [14, 15]. It measures the average similarity between each data cluster. Lower DBI values indicates better class separation [28]. Note that in our case there is no need for clustering per se as the clusters are simply the classes (idle vs. task states) we wish to detect. In this paper, applying artefact handling methods (***BSS-CCA, ICA and PCA***) to all channels will be compared with applying these methods to the selected EMG-CCh only.

The extracted features were sorted in ascending order based on DBI values. The features that had smaller DBI values (i.e., better class separation) were used for further evaluation. In BCIs, as in most human-machine interaction systems, the minimum number of features is recommended to reduce computational loads. Thus, three sets based on a) the smallest DBI, b) smallest 1%, and c) smallest 5% were used for evaluation.

Our EMG-CChs method and typical artefact handling methods were compared using the DBI values. The Wilcoxon test for significance and the statistical power t-test were applied to the DBI values.

H. Reliability of the EMG-CCh Selection Method

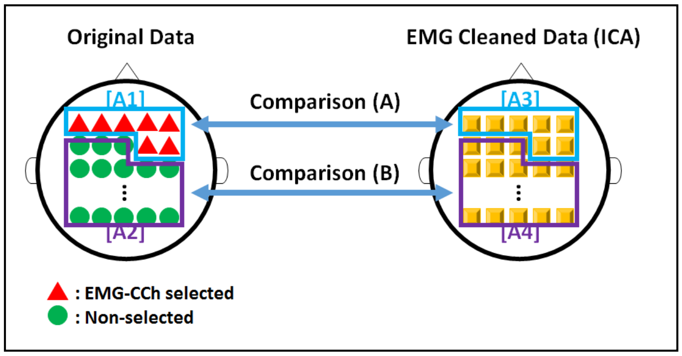

Figure 7. Example reliability test of the EMG-CCh selection method.

Two questions must be asked concerning the proposed EMG-CCh selection method. First, is it actually selecting the EMG-contaminated channels that have class-dependent EMG data? And, secondly, is it strict enough to be used? In order to ensure its reliability, we have gone through extra testing by statistically comparing ICA-processed EMG-free data with our suggested method using both our data and the BCI competition data set. The testing procedure is as follows: Figure 7 shows how to separate channel data for the test. The testing method will be explained based on the example figure. Firstly, ICA was applied to all channels to remove EMG artefacts. Thus, it can be confirmed that right side of the figure has EMG-free signals. Then, on the left panel, channels were separated based on the EMG-CCh selection method (area ***A1*** for EMG-CCh, and ***A2*** for artefact-free channels). Based on this, the EMG-free data on the right panel is also separated into ***A3*** and ***A4*** using the channel numbers from the left panel. Secondly, feature extraction with an autoregressive model was processed in each area. Then, the data with the smallest 10% of DBI values between idle and task state were selected from sets ***A1*** to ***A4***. Finally, a Wilcoxon test was applied between the best DBIs from ***A1*** and ***A3*** (**Comparison (A)**) as well as between the DBIs from ***A2*** and ***A4*** (**Comparison (B)**). In an ideal EMG-CCh selection scenario, **Comparison (A)** should show significant difference between ***A1*** and ***A3*** as ***A1*** contains EMG artefacts whereas ***A3*** is clean, so the DBI value from ***A1*** shows significantly lower DBI values, which indicates that EMG artefacts would have played a significant role in class separation.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2018.2847316, IEEE Transactions on Neural Systems and Rehabilitation Engineering

Also, **Comparison (B)** should show no significant DBI difference between ***A2*** and ***A4***. Based on the EMG-CCh selection method, ***A2*** channels were found not to have an EMG-related role in class separation. Thus, it should give similar (i.e, no significant difference) DBI values to those from set ***A4***. Therefore, if **Comparison (A)** shows significant difference and **Comparison (B)** shows no significant difference, it can be concluded that the suggested EMG-CCh selection method correctly chose EMG artefacts contaminated channels that would otherwise have played an EMG-contaminated role in class separation and the remaining channels did not play any significant class-dependent result.

III. Results and Discussion

A. Sound-Production Related Cognitive Tasks Onset Detection Data

Table I shows an example from participant 1 with the inhibited overt siren (IO_Siren) task onset detection. The smallest 5% of DBI values (384 features x 5% = 19 values) were selected from the AR features. The Wilcoxon test *p*-values and statistical power *t*-test were calculated for ***No EMG handling vs. EMG-CCh removal, ICA vs. ICA (EMG-CCh), PCA vs. PCA (EMG-CCh) and BSS-CCA vs. BSS-CCA (EMG-CCh),*** respectively, with the 19 DBI values. The results show that applying EMG removal methods only to the EMG-CCh significantly improves class separation between Idle and Task states (*p*-value < 0.05) when using ICA (EMG-CCh) and BSS-CCA (EMG-CCh), and the power is statistically conclusive (*t*-test > 0.5). These evaluations were done with all seven subjects and with all four different onset detection tasks. In addition, the single smallest, smallest 1%, and smallest 5% DBI results, respectively were separately evaluated with the four different feature extraction methods.

Figure 8 shows DBI comparison results. 5% of the smallest DBI values were taken from each feature domain: 19 (384 AR model features x 0.05), 22 (448 Band power features x 0.05), 16 (320 DWT features x 0.05) DBI values respectively. The blue areas in the figure indicate that applying EMG handling methods only to the EMG-CCh (our proposed method) shows significant improvement (Wilcoxon test *p*-value < 0.05) in class separation, and power is statistically conclusive (statistical power *t*-test > 0.5). On the other hand, red areas represent instances in which our proposed method showed significantly higher DBIs and, thus, less class separation between Idle and Task states. The grey horizontal stripe areas showed no significant difference between our method and typical EMG handling techniques when applied to all channels.

Table I. Example of Wilcoxon test and Statistical power calculations for features with the lowest 5% DBI values.

Average of Smallest 5% DBI values				
AR Model Feature [P1 – Inhibited Overt Siren Task]	No EMG handling [std]	ICA [std]	PCA [std]	BSS-CCA [std]
	4.090 [±0.44]	6.756 [±0.66]	4.179 [±0.59]	4.227 [±0.27]
	EMG-CCh removal [std]	ICA(EMG-CCh) [std]	PCA(EMG-CCh) [std]	BSS-CCA(EMG-CCh) [std]
	4.096 [±0.45]	4.096 [±0.45]	4.086 [±0.44]	3.978 [±0.40]
Wilcoxon-test (*p*-values)	-0.9069	≈ 0	0.1568	0.0296
Statistical Power (*t*-test)	0.0502	≈ 1	0.0758	0.7863

The Wilcoxon-test p-value is shown as a negative number if standard EMG handling processes [ICA / PCA / BSS-CCA] give smaller average DBI value than [ICA (EMG-CCh) / PCA (EMG-CCh) / BSS-CCA (EMG-CCh)].

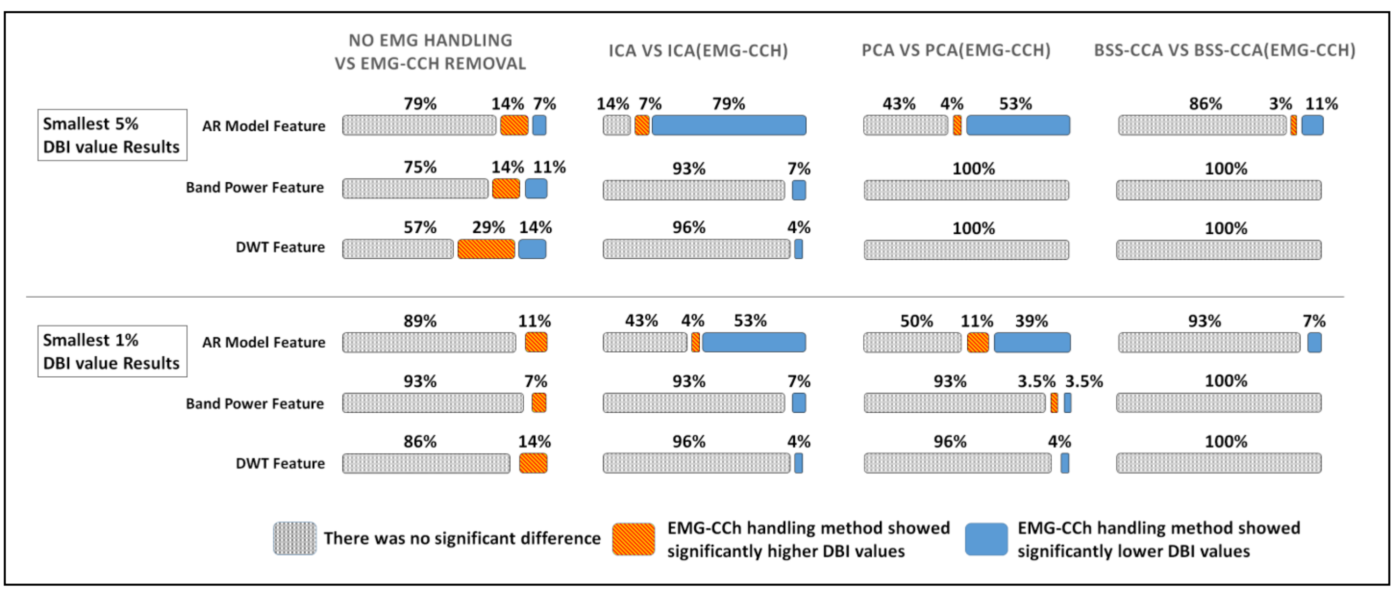

Figure 8. Results with features with the smallest 5% and 1% DBI values using the onset detection data set.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2018.2847316, IEEE Transactions on Neural Systems and Rehabilitation Engineering

Table II. Results with the features that gave the smallest DBI value, onset detection data set.

AR Model Feature [Average of 7 subjects & 4 onset Detection Tasks]	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	3.569 [±1.72]	4.351 [±1.86]	4.119 [±2.09]	3.713 [±1.85]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	3.320 [±2.01]	3.769 [±1.87]	3.648 [±1.76]	3.547 [±1.70]
Band Power Feature [Average of 7 subjects & 4 onset Detection Tasks]	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	2.494 [±1.15]	2.566 [±1.15]	2.592 [±1.20]	2.500 [±1.15]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	2.590 [±1.21]	2.506 [±1.18]	2.514 [±1.19]	2.496 [±1.15]
DWT Feature [Average of 7 subjects & 4 onset Detection Tasks]	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	2.741 [±1.26]	2.825 [±1.29]	2.740 [±1.20]	2.753 [±1.25]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	2.946 [±1.39]	2.777 [±1.23]	2.762 [±1.27]	2.745 [±1.24]
CSP Feature [Average of 7 subjects & 4 onset Detection Tasks]	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	3.794 [±7.14]	4.724 [±9.85]	5.977 [±16.10]	3.810 [±7.22]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	3.828 [±6.21]	3.839 [±7.08]	8.320 [±25.52]	3.767 [±7.11]

As can be seen from the case of ***No EMG handling vs. EMG-CCh removal***, there is no significant difference most of the time. But, average DBI values become higher more times than become lower for all three different feature extraction methods. This an expected result as EMG artefacts could play a role in class separation if they are not handled properly. Thus, removing EMG-CCh reduces class separation, but it shows why artefact handling is required. In the case of ***ICA vs. ICA (EMG-CCh)***, 79%, 7% and 4% out of 28 tests (7 subjects x 4 onset tasks) showed significant class separation improvements with EMG-CCh selection for AR model features, band power and DWT features, respectively. On the other hand, only 7% of the AR features gave higher DBI values with the proposed EMG-CCh method and all remaining tests yielded no significant difference. In the ***PCA vs. PCA (EMG-CCh)*** case, band power and DWT feature made no significant difference but AR features showed significant class separation improvement for 53% of the tests and 43% remained as not significantly different. In the ***BSS-CCA vs. BSS-CCA (EMG-CCh)*** case, 11% showed significant class separation improvement with AR features, while 3% became worse and most tests (86%) showed no significant difference. Band power and DWT features also showed no significant difference in class separation between BSS-CCA applied to all channels vs. only to EMG-CCh.

The smallest 1% DBI results showed similar trend as results based on the best 5% DBIs. While results generally show no significant difference between our method and standard EMG-removal techniques, EMG-CCh selection did yield significantly better class separation more often than it yielded worse separation.

Table II lists the (single) smallest DBI value in each case. CSP features were included only in this table because it has just one feature point. Also, as this table shows only the smallest DBI value, Wilcoxon tests and Statistical power *t*-tests could not be applied. However, even though no inferences can be made concerning statistical significance on a test-by-test basis, the smallest DBI values can give relevant information that might be useful to other BCI studies as minimising the number of data points is always an important goal in BCIs. For this reason, we averaged the smallest DBI values from 28 tests and compared overall results between our EMG-CCh method and standard EMG handling methods (there was no statistical difference $p > 0.05$ from all cases). The averaged smallest DBI values were lower with our EMG-CCh handling method in most cases, except for DWT features and for CSP features with PCA. In terms of CSP featured, some subjects had huge DBI values (i.e., poor class separation) compared to other feature domains. But, in general, our EMG-CCh method gave lower DBI values.

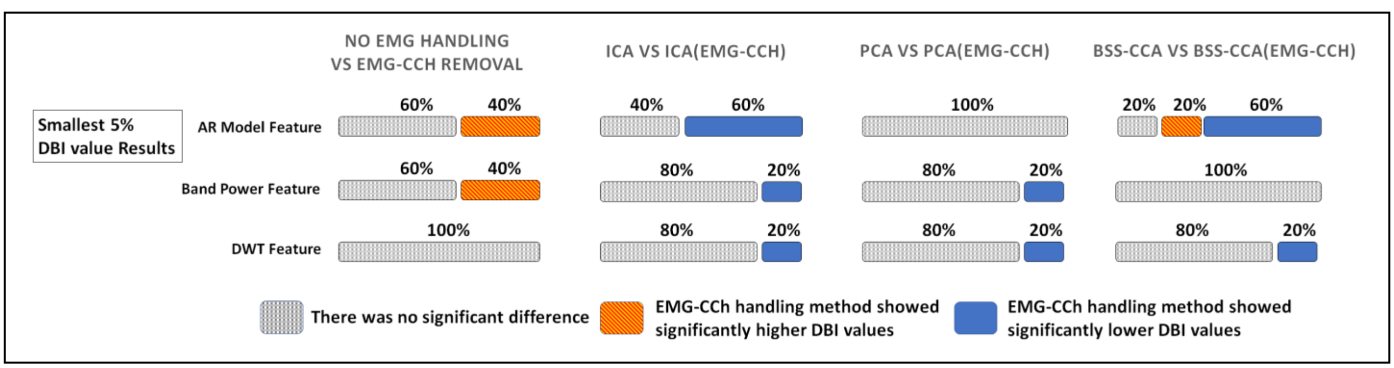

Figure 9. Result of smallest 5% DBI values from BCI competition data set.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSRE.2018.2847316, IEEE Transactions on Neural Systems and Rehabilitation Engineering

Table III. Result of the smallest DBI values from BCI competition data set

AR Model Feature [Average of 5 subjects]	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	8.623 [±3.23]	11.689 [±4.64]	9.551 [±4.89]	10.753 [±3.13]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	11.393 [±6.19]	10.362 [±5.34]	9.846 [±5.23]	9.768 [±3.95]
Band Power Feature [Average of 5 subjects]	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	4.968 [±1.50]	5.492 [±1.89]	5.101 [±1.22]	9.244 [±2.91]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	5.175 [±1.82]	5.175 [±1.82]	4.963 [±1.45]	5.239 [±1.89]
DWT Feature [Average of 5 subjects]	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	6.756 [±2.52]	7.192 [±2.90]	6.984 [±2.32]	12.408 [±2.95]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	6.957 [±2.84]	6.951 [±2.83]	6.780 [±2.55]	6.940 [±2.70]
CSP Feature [Average of 5 subjects]	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	3.467 [±1.72]	3.722 [±2.03]	5.387 [±4.05]	4.637 [±2.60]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	4.292 [±2.17]	3.806 [±1.88]	3.791 [±1.34]	3.281 [±1.44]

B. BCI Competition Data Set

The smallest 5% DBIs gave 7, 8 and 6 features from the AR model, band power and DWT domains, respectively. There was only one feature when using the smallest 1% DBIs. Thus, only these two cases were analysed with the BCI competition data set.

In Figure 9, 40% of participants showed less class separation with EMG-CCh removal using AR model and band power features in the case of comparisons between ***No EMG handling vs. EMG-CCh removal***. The remaining 60% has no statistical difference in the latter comparison. As before, this is an expected result as EMG would have had a role in class separation.

In the case of ***ICA vs. ICA (EMG-CCh)***, 60%, 20% and 20% of 5 subjects - for each of the three feature domains, respectively - showed significantly lower DBI values with our EMG-CCh method and none of participants yielded worse class separation with our method. Comparing ***PCA vs. PCA (EMG-CCh)***, only one participant out of five yielded significantly improved class separation with our method. In the ***BSS-CCA vs. BSS-CCA (EMG-CCh)*** similar results were seen for band power and DWT features. For AR features, on the other hand, 60% (3 subjects) showed significant improvement while 1 participant had higher DBI value with the EMG-CCh method. Table III shows a similar trend; it mostly has lower DBI values with the EMG-CCh method except for ICA (EMG-CCh) with CSP feature and PCA (EMG-CCh) with AR model feature.

C. Reliability of the EMG-CCh Selection Method

As explained in section *II.H* above, we applied two types of comparisons, A and B, respectively, to test for the possibility that we did not eliminate all EEG channels with significant class-dependent EMG contamination.

Using our sound-production onset data and the average of the smallest 10% DBI values from 28 cases (7 subjects x 4 tasks), **Comparison A** showed significant difference (p=0.011, mean value of ***A1=6.26,*** mean value of ***A3=8.11***) while **Comparison B** showed there was no significant difference (p=0.124, mean value of ***A2=4.85, A4=5.00***). The results were similar when we used the BCI competition data set (5 subjects). In this case results also showed a significant difference in **Comparison A** (p=0.016, mean value of ***A1=12.61, A3=25.88***) and no significant difference in **Comparison B** (p=0.222, mean of ***A2=18.27, A4=20.94***). From these results, it can be said that our EMG-CCh selection method is indeed correctly identifying channels that could affect class-dependent results.

IV. Conclusions

EMG artefact handling is an essential procedure for EEG based studies (e.g., BCIs, brain mapping and clinical tests). However, common BSS techniques can cause loss of useful EEG information [6-9]. For this reason, we proposed a new technique for selecting EEG channels contaminated with class-dependent EMG artefacts (called EMG-CCh) to minimise information loss and improve class separation.

In order to ensure the method is actually selecting EEG channels that have class-dependent EMG data and that the method is strict enough to be used, a statistical reliability test (comparing with ICA results) had been done. The test results showed that the proposed EMG-CCh selection method is indeed correctly identifying channels that could affect class-dependent results.

With autoregressive model features, 79% (ICA), 53% (PCA) and 11% (BSS-CCA) of 28 tests showed significant improvement (Wilcoxon $p < 0.05$; statistical power > 0.5) using our data with the proposed method. With BCI competition data, we observed that 60% (ICA) and 60% (BSS-CCA) out of 5 subjects yielded significantly better with the EMG-CCh method.

In summary, the proposed method showed significant class separation improvement (compared to existing techniques) with both our data and the BCI competition data set in many cases. Also, the method can be used on its own or it can be combined with pre-existing artefact handling methods.

References

[1] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211-1279, 2012.

[2] J. S. Barlow, "Artifact processing (rejection and minimization) in EEG data processing," *Handbook of electroencephalography and clinical neurophysiology*, vol. 2, pp. 15-62, 1986.

[3] J. S. Barlow, "EMG artifact minimization during clinical EEG recordings by special analog filtering," *Electroencephalography and clinical neurophysiology,* vol. 58, no. 2, pp. 161-174, 1984.

[4] K. L. Coburn and M. A. Moreno, "Facts and artifacts in brain electrical activity mapping," *Brain Topography,* vol. 1, no. 1, pp. 37-45, 1988.

[5] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey," *Clinical neurophysiology,* vol. 118, no. 3, pp. 480-494, 2007.

[6] A. Vergult *et al.*, "Improving the interpretation of ictal scalp EEG: BSS–CCA algorithm for muscle artifact removal," *Epilepsia,* vol. 48, no. 5, pp. 950-958, 2007.

[7] H. Nam, T. G. Yim, S. K. Han, J. B. Oh, and S. K. Lee, "Independent component analysis of ictal EEG in medial temporal lobe epilepsy," *Epilepsia,* vol. 43, no. 2, pp. 160-164, 2002.

[8] E. Urrestarazu, J. Iriarte, M. Alegre, M. Valencia, C. Viteri, and J. Artieda, "Independent component analysis removing artifacts in ictal recordings," *Epilepsia,* vol. 45, no. 9, pp. 1071-1078, 2004.

[9] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Advances in neural information processing systems,* pp. 145-151, 1996.

[10] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *IEEE transactions on Biomedical Engineering,* vol. 53, no. 12, pp. 2583-2587, 2006.

[11] Y. Song and F. Sepulveda, "Classifying speech related vs. idle state towards onset detection in brain-computer interfaces overt, inhibited overt, and covert speech sound production vs. idle state," in *Biomedical Circuits and Systems Conference (BioCAS), 2014 IEEE*, 2014, pp. 568-571: IEEE.

[12] Y. Song and F. Sepulveda, "A novel onset detection technique for brain–computer interfaces using sound-production related cognitive tasks in simulated-online system," *Journal of Neural Engineering,* vol. 14, no. 1, p. 016019, 2017.

[13] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI Competition 2008–Graz data set A," *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology,* p. 16, 2008.

[14] J. C. Bezdek and N. R. Pal, "Some new indexes of cluster validity," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on,* vol. 28, no. 3, pp. 301-315, 1998.

[15] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* no. 2, pp. 224-227, 1979.

[16] I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "EMG contamination of EEG: spectral and topographical characteristics," *Clinical Neurophysiology,* vol. 114, no. 9, pp. 1580-1593, 2003.

[17] V. Krishnaveni, S. Jayaraman, S. Aravind, V. Hariharasudhan, and K. Ramadoss, "Automatic identification and Removal of ocular artifacts from EEG using Wavelet transform," *Measurement science review,* vol. 6, no. 4, pp. 45-57, 2006.

[18] M. Kirkove, C. François, and J. Verly, "Comparative evaluation of existing and new methods for correcting ocular artifacts in electroencephalographic recordings," *Signal Processing,* vol. 98, pp. 102-120, 2014.

[19] J. Gao, C. Zheng, and P. Wang, "Online removal of muscle artifact from electroencephalogram signals based on canonical correlation analysis," *Clinical EEG and neuroscience,* vol. 41, no. 1, pp. 53-59, 2010.

[20] G. Barbati, C. Porcaro, F. Zappasodi, P. M. Rossini, and F. Tecchio, "Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals," *Clinical Neurophysiology,* vol. 115, no. 5, pp. 1220-1232, 2004.

[21] A. Greco, N. Mammone, F. C. Morabito, and M. Versaci, "Kurtosis, Renyi's entropy and independent component scalp maps for the automatic artifact rejection from EEG data," *International Journal of Signal Processing,* vol. 2, no. 4, pp. 240-244, 2006.

[22] P. Sharmilakanna and R. Palaniappan, "EEG artifact reduction in VEP using 2-stage PCA and N4 analysis of alcoholics," in *Intelligent Sensing and Information Processing, 2005. ICISIP 2005. Third International Conference on*, 2005, pp. 1-7: IEEE.

[23] X. Chen, C. He, and H. Peng, "Removal of muscle artifacts from single-channel EEG based on ensemble empirical mode decomposition and multiset canonical correlation analysis," *Journal of Applied Mathematics,* vol. 2014, 2014.

[24] J. P. Burg, "A new analysis technique for time series data," 1968.

[25] K. Brigham and B. Kumar, "Imagined speech classification with EEG signals for silent communication: a preliminary investigation into synthetic telepathy," in *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on*, 2010, pp. 1-4: IEEE.

[26] Y. Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selelction in motor imagery based brain-computer interface," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2006, pp. 5392-5395: IEEE.

[27] Y. J. Song and F. Sepulveda, "Classifying siren-sound mental rehearsal and covert production vs. idle state towards onset detection in brain-computer interfaces," in *Brain-Computer Interface (BCI), 2015 3rd International Winter Conference on*, 2015, pp. 1-4: IEEE.

[28] F. Kovács, C. Legány, and A. Babos, "Cluster validity measurement techniques," in *Proceedings of the 6th International Symposium of Hungarian Researchers on Computational Intelligence, Budapest*, 2005, pp. 18-19: Citeseer.