

2015-03

Automated identification of neural correlates of continuous variables

Daly, I

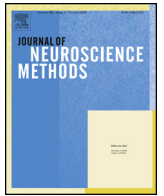
<http://hdl.handle.net/10026.1/6523>

10.1016/j.jneumeth.2014.12.012

Journal of Neuroscience Methods

Elsevier BV

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.



Computational Neuroscience

Automated identification of neural correlates of continuous variables



Ian Daly^{a,*}, Faustina Hwang^a, Alexis Kirke^b, Asad Malik^a, James Weaver^a,
Duncan Williams^b, Eduardo Miranda^b, Slawomir J. Nasuto^a

^a Brain Embodiment Lab, School of Systems Engineering, University of Reading, Reading, UK

^b Interdisciplinary Centre for Computer Music Research, Plymouth University, Plymouth, UK

HIGHLIGHTS

- A new method for identifying EEG correlates of continuous independent variables.
- Our method outperforms canonical correlation analysis and common spatial patterns.
- When applied to real EEG during music it finds known correlates of music tempo.
- The method also identifies novel neural correlates of music induced emotion.

ARTICLE INFO

Article history:

Received 29 October 2014

Received in revised form

16 December 2014

Accepted 17 December 2014

Available online 26 December 2014

Keywords:

Feature selection

Eigen-decomposition

Neural correlates

Electroencephalogram (EEG)

ABSTRACT

Background: The electroencephalogram (EEG) may be described by a large number of different feature types and automated feature selection methods are needed in order to reliably identify features which correlate with continuous independent variables.

New method: A method is presented for the automated identification of features that differentiate two or more groups in neurological datasets based upon a spectral decomposition of the feature set. Furthermore, the method is able to identify features that relate to continuous independent variables.

Results: The proposed method is first evaluated on synthetic EEG datasets and observed to reliably identify the correct features. The method is then applied to EEG recorded during a music listening task and is observed to automatically identify neural correlates of music tempo changes similar to neural correlates identified in a previous study. Finally, the method is applied to identify neural correlates of music-induced affective states. The identified neural correlates reside primarily over the frontal cortex and are consistent with widely reported neural correlates of emotions.

Comparison with existing methods: The proposed method is compared to the state-of-the-art methods of canonical correlation analysis and common spatial patterns, in order to identify features differentiating synthetic event-related potentials of different amplitudes and is observed to exhibit greater performance as the number of unique groups in the dataset increases.

Conclusions: The proposed method is able to identify neural correlates of continuous variables in EEG datasets and is shown to outperform canonical correlation analysis and common spatial patterns.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The electroencephalogram (EEG) is a method for measuring changes in electro-potential in the cortex related to the activation levels of cortical neuronal populations (Niedermeyer and Silva, 2005). It is a popular method for studying neuro-electrophysiological correlates of cognitive processes and behaviour.

The EEG has a high temporal resolution and a wide spectral range and may, therefore, be described by a very large number of features. These include, for example, the band power within specific frequency ranges, amplitudes over specific regions of the cortex, or measures of interactions between different spatial regions (Rahman et al., 2012). However, EEG has very poor signal to noise ratio (SNR), which often means that many repeated trials are required before cognitively relevant information emerges from the background noise present in the signal (Niedermeyer and Silva, 2005).

Due to practical limits on the numbers of repetitions of cognitive events participants in cognitive experiments may perform, an

* Corresponding author. Tel.: +44 1183788609.
E-mail address: i.daly@reading.ac.uk (I. Daly).

investigative researcher is often faced with a very large potential feature space and a very small number of trials. Thus, identification of reliable task-related features is a considerable challenge.

A number of approaches may be taken to tackle this problem. Where particular cognitive processes within the EEG are phase-locked to the trial commencement time and stationary, a time-averaging approach may be adopted to attempt to identify features related to specific neural correlates. However, in cases where physiological responses are not phase-locked (for example band-power measures [Pfurtscheller and Lopes da Silva, 1999](#)) an approach from machine learning may be adopted ([Alpaydin, 2004](#)).

One such approach which has gained considerable traction in recent years is common spatial patterns (CSP), which is based upon eigen-decomposition of the covariance matrices of each group in the dataset ([Koles et al., 1990](#)). This method was used originally in the brain-computer interface (BCI) ([Wolpaw et al., 2002](#)) community to identify optimal features for separating two groups of events in the EEG and has since gained growing popularity for a range of uses, for example, identifying neural activity related to motor imagery ([Friedrich et al., 2012](#)). Extensions of the method have also been proposed for multiple group cases ([Grosse-Wentrup and Buss, 2008](#)).

However, while CSP is able to identify features which may be applied to optimally separate discrete groups of tasks, it is less effective in the case of continuous variables. Thus, in the case of correlation studies, where neural correlates are sought relating to a continuous independent variable, it may not be the most suitable method.

An alternative approach, which may be applied in this case is canonical correlation analysis (CCA) ([Knapp, 1978](#)). CCA attempts to find relationships between sets of independent variables, for example between two or more time series, and may be used to identify neural correlates of continuous variables ([Hardoon et al., 2004](#)). However, CCA is only able to identify sets of variables which linearly correlate with the independent variables and, therefore, may be limited in its applications ([Hardoon et al., 2004](#)).

To tackle these problems with CSP and CCA we have developed an alternative automated feature selection method that is able to identify neural correlates of continuous independent variables. The proposed method is based upon eigen-decomposition of the coarse-grained (rescaled) combined matrix of features and the continuous independent variable.

We first describe the method. We then compare the method to CSP and CCA on a synthetic test dataset before using it to attempt to find feature sets which correspond to a continuous independent variable in a study of neural correlates of music-induced emotions and neural correlates of music perception.

2. Methods

2.1. Proposed method

The proposed method aims to identify features which covary with an independent variable by first redistributing, then coarse-graining, performing principal component analysis (PCA), and clustering the set of available features. This can be summarized as follows.

1. Prepare the data set by first uniformly redistributing the values each feature takes.
2. Then coarse-grain the values each feature takes across all trials.
3. Calculate a covariance matrix from the coarse-grained feature matrix to measure relationships within the dataset.
4. Perform principal component analysis (PCA) to identify directions of maximum variance.

5. Calculate the participation index, sort and extract the top q index values. These index the features of interest.

For a data set $\mathbf{D} \in \mathbb{R}^{M \times N}$ of N trials, each trial may be represented by M features. Features may describe the data in a number of different ways, for example, representing them in the time and/or frequency domains. A feature vector comprises the values of that feature for all trials in the dataset \mathbf{D} . A feature vector is defined as

$$\mathbf{F}_i = [D_{i,1}, \dots, D_{i,N}], i \in [1, \dots, M], \quad (1)$$

where i denotes the i 'th feature in the data set. An additional vector is concatenated to the set of feature vectors, resulting in a total of $M+1$ feature vectors. This additional vector will contain the values taken by the continuous independent variable across the N trials. We will refer to this additional variable as the independent covariate vector.

In order to improve the robustness of the selection of features related to the independent covariate vector, the range of values in the set of all feature vectors is uniformly redistributed over the range $1, \dots, u$, where u denotes the number of unique values in the independent covariate vector, and coarse-grained into u partitions. This is done by first z-scoring the coordinates of each feature vector to a mean of zero and a standard deviation of 1.

To improve the robustness of the feature selection in subsequent steps, the complementary error function ($1 - \text{the error function}$) is then applied to the normalised feature vectors ([Hunter and Regan, 1972](#)). Consequently, this allows us to use PCA to identify features in the data. Finally, the coarse graining is completed by affine-mapping the entries in each feature vector so that they fall into the range $1, \dots, u$ and rounding the values in the feature vector in order to ensure that the number of discrete values taken by the members of each feature vector is equal to u . This is done by dividing the values within each feature vector by the maximum value in the feature vector, multiplying them by u , and adding 1.

This results in a set of coarse-grained feature vectors $\bar{\mathbf{F}}_i = [\bar{D}_{i,1}, \dots, \bar{D}_{i,N}]$, where $\bar{D}_{i,n}, n \in 1, \dots, N, i \in 1, \dots, (M+1)$ denotes the coarse-grained value of feature i , trial n .

It is now possible to identify informative features by employing techniques related to finding clusters in multivariate datasets. One of the clusters obtained as a result will contain the independent covariate vector. The elements of this grouping therefore define the features that are most closely related to it. Thus, these features are taken to be the features that optimally relate to the independent covariate vector.

To this end we adapted a form of spectral clustering introduced in ([Allefeld et al., 2007](#)). The $(M+1) \times (M+1)$ covariance matrix Σ is calculated from the coarse-grained collection of feature vectors. Eigen decomposition is applied to the covariance matrix Σ to find a set of eigenvalues and eigenvectors defined as

$$\Sigma \mathbf{V}_k = \lambda_k \mathbf{V}_k \quad (2)$$

where λ_k denotes the eigenvalues and \mathbf{V}_k the eigenvectors.

Eigenvalues are then sorted in descending order and the q eigenvalues in the top 5th percentile are identified. The corresponding eigenvectors explain the majority of the variance in the set of feature vector projections onto the linear subspace spanned by these eigenvectors $\mathbf{V} = [v_{i,k}], k \in [1, \dots, q], \forall i \in [1, \dots, (M+1)]$.

The Participation Index (\mathbf{PI}) (defined in ([Allefeld et al., 2007](#))) provides a measure of the involvement of each feature vector in each cluster

$$PI_{i,k} = \lambda_k * v_{i,k}^2, \forall i = [1, \dots, (M+1)], \forall k = [1, \dots, q], \quad (3)$$

where the eigenvalues λ_k and their corresponding eigenvectors have been pre-sorted in descending order of eigenvalue.

We first inspect the column of \mathbf{PI} corresponding to the independent covariate vector ($PI_{(M+1),:}$). The largest \mathbf{PI} in this column

indexes an eigenvector indicating the cluster of feature vectors most involved with the independent covariate vector. Thus, the corresponding feature vectors in this cluster represent a good set of features that relate to the independent covariate vector.

Therefore, the final set of selected features are identified by selecting the top q \mathbf{PI} s from the row $PI_{:, \alpha}$, where α denotes the index of the largest \mathbf{PI} in the column $PI_{(M+1), :}$. These indexes identify the optimal set of features relating to the independent covariate vector and, hence, the set of identified features.

2.2. Common spatial patterns

Common spatial patterns attempt to identify a spatial filter that maximally separates data between two or more groups (Koles et al., 1990). They have been used for many applications, primarily in brain–computer interfacing (BCI) research (Hwang et al., 2013).

For sets of features $X_1 \in \mathbb{R}^{t_2 \times n}$ and $X_2 \in \mathbb{R}^{t_2 \times n}$ corresponding to different conditions, where n is the number of samples and t_1 and t_2 denote the number of trials in each group, CSP attempts to find a filter ($w \in \mathbb{R}^{n \times n}$)^T such that the variance ratio between the two groups is maximised. This is defined as

$$w = \underset{w}{\operatorname{argmax}} \frac{\|X_1 w\|^2}{\|X_2 w\|^2}. \quad (4)$$

The solution is found by solving the generalised eigenvalue problem on w . The eigen-decomposition is used to identify the eigenvectors and eigenvalues of the covariance matrices. The largest eigenvector may then be used as a filter, denoting elements of the dataset that separate the groups.

2.3. Canonical correlation analysis

Canonical correlation analysis (CCA) attempts to find linear combinations of two vectors of random variables such that they maximally correlate with one another. For example, in the case where one of the vectors contains just one random variable and the other several random variables, the single random variable may denote an independent covariate and CCA could be used to identify the set of random variables which maximally correlate with this independent covariate (Hardle and Simar, 2007).

Given two column vectors of random variables $X = (x_1, \dots, x_n)^T$ and $Y = (y_1, \dots, y_n)^T$, a covariance matrix is defined as $\Sigma_{XY} = \operatorname{cov}(X, Y)$. CCA then seeks two vectors a and b for which the correlation function is $p = \operatorname{corr}(a^T X, b^T Y)$. These vectors then give weightings to the relative correlation of each random variable in X and Y . Hence, they may be used to select optimal sub-sets of random variables in either X and/or Y (Hardoon et al., 2004).

2.4. Comparison

Artificial event-related potentials (ERPs) were used to compare our proposed method to CSP and CCA. An ERP is a phase-locked change in amplitude of the EEG that occurs in response to a stimulus or event. ERPs may be used to study a wide variety of cognitive processes and are amongst the most studied phenomena in the EEG (Handy, 2005).

For the synthetic dataset, we simulated an EEG component whose amplitude correlated with an independent covariate, where the number of values that the independent covariate could take was varied. Thus, we sought to evaluate the performance of the methods for datasets containing trials with n different amplitudes in the ERPs. The amplitude of the ERPs correlated with the value of the independent covariate vector.

EEG data was generated via a neural mass model (David and Friston, 2003) simulating 100 trials recorded from 19 EEG electrode recording channels of length 1 s each at a sample rate of 125 Hz.

The dataset of 100 trials was split into groups of trials with equal numbers of trials per group, and the number of groups was varied between 2 and 20 (e.g. in the case of 2 groups, each group contained 50 trials). In the case that of 100 trials not evenly splitting into the desired number of groups, uniformly drawn trials were removed from the dataset until an even split of trials across groups was possible.

ERPs were simulated by increasing the amplitude of the synthetic EEG between 0.2 s and 0.8 s by a gain-adjusted Gaussian function. The gain increased the amplitude of the ERPs as a function of the value of the independent covariate vector. Thus, synthetic ERPs added to group 10 were larger in amplitude than synthetic ERPs added to group 9. Thus, the added ERP component was a function of the independent covariate and amenable to identification via the proposed method, CSP, and CCA.

ERPs were added to channels 14, 15, and 16 (channels corresponding to positions P3, Pz, and P4 in the case of a real EEG recorded via the International 10/20 system for electrode placement). Datasets were generated with numbers of groups of trials ranging between 2–20 and the performance of

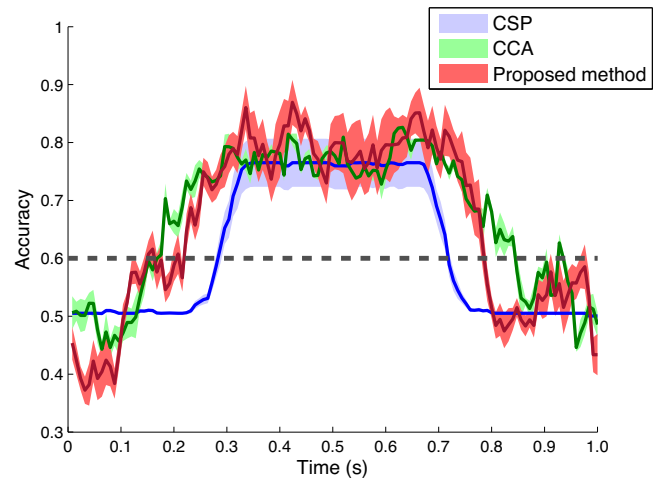


Fig. 1. Classification accuracy over time for artificial ERPs (2 groups) with features selected via our method, common spatial patterns (CSP), and canonical correlation analysis (CCA). The dashed line indicates the level of statistically significant classification accuracy ($p < 0.05$).

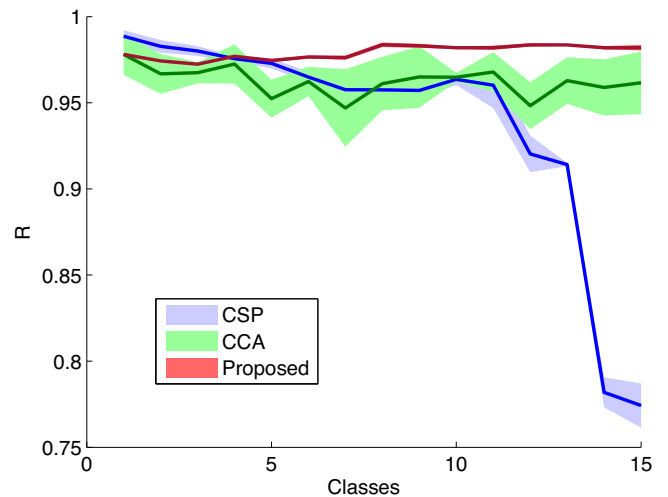


Fig. 2. Correlation between the features selected by each of the methods and the values of the independent covariate vector as a function of the number of unique values in the independent covariate. The shaded lines represent ± 2 standard deviations.

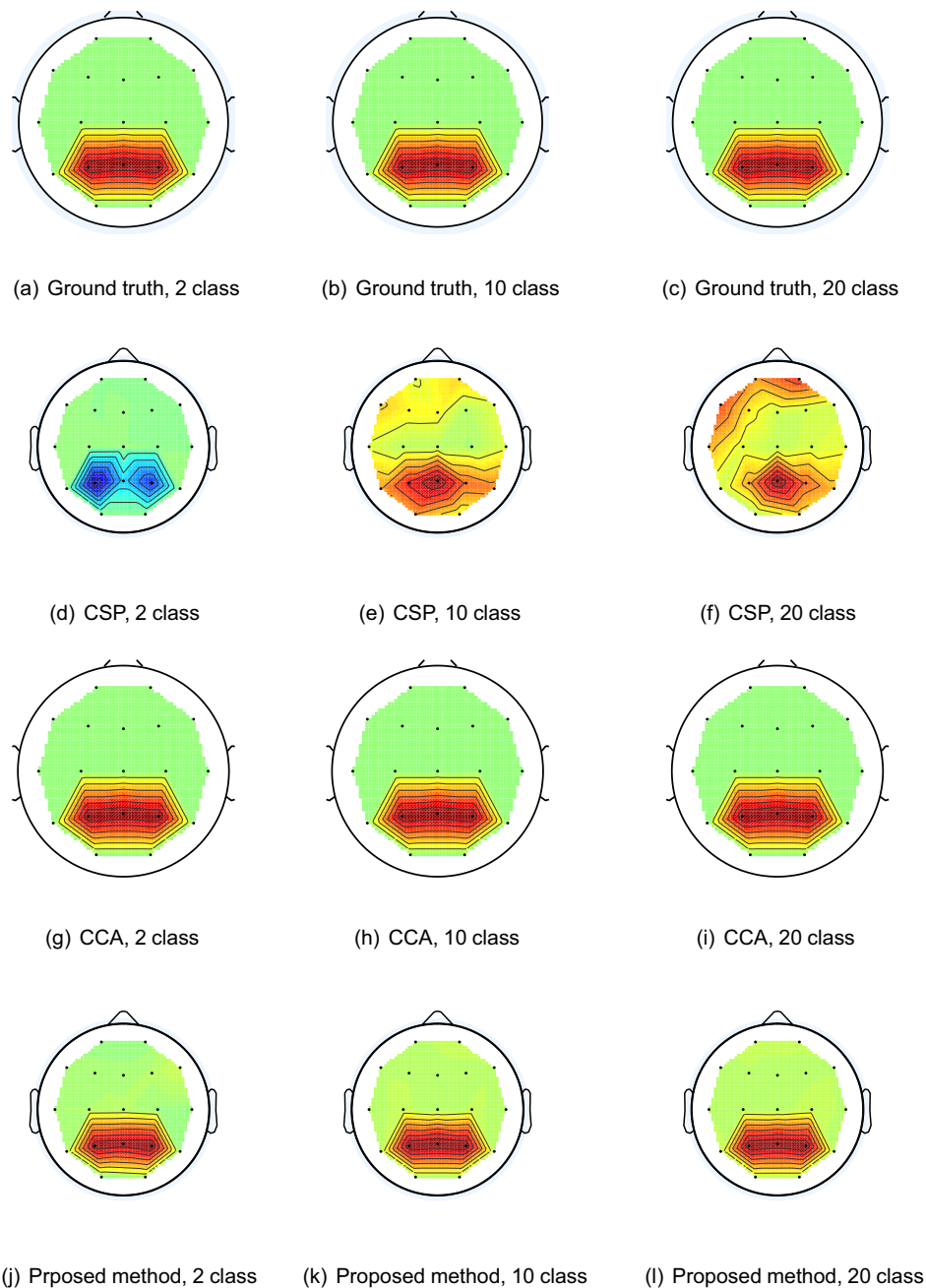


Fig. 3. Topographic scalp maps of channels selected via the CSP and CCA feature selection methods and our proposed method for 2, 10, and 20 groups (classes) of ERPs indicated by unique values in the independent covariate vector. Ground truth maps illustrate which channels have had simulated ERPs added.

each of the feature selection methods was evaluated for each dataset.

In the case of two groups of EEG trials, performance was measured via a linear discriminant analysis (LDA) classifier, which was used to attempt to differentiate trials via the amplitude of the identified features. In the case of more than 2 unique values in the independent covariate the correlation between the filtered (selected) EEG features and the independent covariate was used to evaluate the performance of the methods. The higher the correlation the better the match between the selected features and the independent covariate.

The method (our proposed method, CSP, or CCA) and the LDA classifier is trained and tested in a 10×10 cross fold train and validation scheme. This allows us to measure the

variance of performance of each of the methods against one another.

2.5. Application

A number of neural correlates of the participant-reported induced emotions and neural correlates of properties of musical stimuli, such as tempo, have been reported elsewhere (Daly et al., 2014, 2014). Here we use our proposed method to first automatically identify some of these neural correlates and second, to search for further, previously unidentified, neural correlates of music-induced affective states.

EEG was first cleaned of artefacts by visual inspection followed by supervised independent component analysis. Further details can

be found in [Daly et al. \(2014\)](#). Band-power features were then extracted from the EEG from a frequency window of width 2 Hz centred at 10 Hz. Time series of the band-power values of length 1 s were extracted from all 19 channels (recorded as per the international 10/20 system) and down-sampled to 100 Hz. Our method was then applied to search for band-power features across all available channels and time points that maximally correlate with each of the following properties:

1. Each of the participant-reported ratings of induced affective states (pleasantness, energy, tension, anger, fear, happiness, sadness, and tenderness).
2. The tempo of the music (measured in beats per minute), estimated via the approach described in [Ellis \(2007\)](#).

3. Results

3.1. Comparison

When considering the two class case, features corresponding to time intervals during which class-dependent EEG has distinct activity are identified by each method, which allows the artificial ERPs to be classified with statistically significant accuracies ($p < 0.01$). Our method achieved a peak accuracy of $0.869 (\pm 0.004)$, while CSP achieved a peak accuracy of $0.765 (\pm 0.021)$, and CCA achieved a peak accuracy of $0.826 (\pm 0.001)$. A 1×3 ANOVA with factor 'Method' and levels 'Proposed', 'CSP', and 'CCA' was used to determine if there is a significant effect on peak accuracy of the method used. A significant effect is found $F(2, 27) = 183.91$, $p < 0.001$. Post-hoc t -tests reveal significant differences between the proposed method and CSP ($p < 0.01$) and the proposed method and CCA ($p < 0.01$). The proposed method has the highest peak accuracy (mean = 0.872 , STD. = 0.005), while CCA (mean = 0.827 , STD. = 0.001) and CSP (mean = 0.756 , STD. = 0.028) have lower peak accuracies.

Additionally, our proposed method achieved a longer period of significant classification accuracy and was detected earlier in the trial. This is illustrated in [Fig. 1](#), in which our proposed method achieves significant classification from 0.2 s to 0.7 s, while CSP only achieves significant accuracy from 0.3 s to 0.7 s, and CCA also achieves significant accuracies from 0.2 s to 0.7 s.

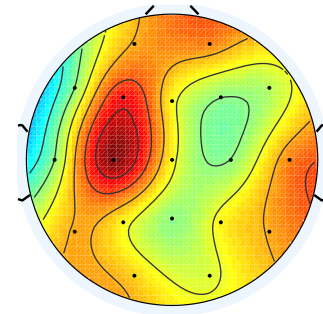
In the case of additional numbers of unique values being added to the independent covariate vector, the ability of each of the feature selection methods to identify good feature sets is measured by the correlation between the selected features and the independent covariate vector. [Fig. 2](#) illustrates how this correlation changes as the number of unique values increases from 2 through to 15.

In the case of between 2 and 5 unique values in the independent covariate vector the performance of the methods is broadly equivalent. However, as the number of unique values increases above 6 the performances of CSP and CCA both exhibit some decline, while our method marginally increases in performance. It is also interesting to note that our proposed method has considerably lower variance than either CCA or CSP, with CSP exhibiting the largest variance in performance.

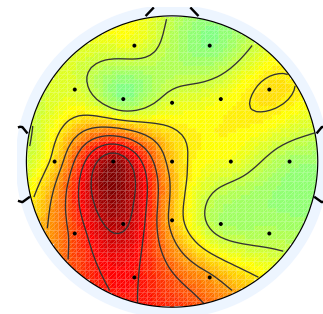
The EEG channels which are selected as optimally separating the groups of trials are illustrated in [Fig. 3](#). Note that as the number of groups increases the precision of the CSP method decreases markedly, while the channels selected by CCA and our proposed method both remain unchanged.

3.2. Application

Our proposed method was also used to identify neural correlates of musical tempo and participant-reported induced affective responses to music. First, neural correlates of tempo were sought



(a) Neural correlates identified by the proposed method



(b) Manually identified correlates

Fig. 4. Neural correlates of musical tempo changes in the 10 Hz frequency band identified by our proposed method (A) and previously manually identified (B) (as described elsewhere in [Daly et al., 2014](#)).

by our method and compared to results previously reported elsewhere ([Daly et al., 2014](#)). [Fig. 4](#) illustrates the mean scalp maps identified as neural correlates of tempo identified by our method within the cross-fold validation method alongside the neural correlates of tempo previously manually identified and reported in [Daly et al. \(2014\)](#).

It is noteworthy that similar patterns of spatial activation are found by our method to those previously reported in [Daly et al. \(2014\)](#). Specifically, left hemisphere motor cortex regions are identified as most strongly relating to the tempo of the music.

Additionally, neural correlates are sought for the participant-reported induced affective states. These are illustrated in [Fig. 5](#), which illustrates the scalp maps identified as corresponding to each of the music-induced emotions.

4. Discussion

The proposed method is able to automatically identify feature sets that relate to continuous independent variables. Therefore, it may be used to identify correlates of continuous variables (for example, neural correlates of music tempo during music listening tasks). Consequently, the proposed method has a wide number of potential applications in correlation studies.

The method is observed to outperform common spatial patterns (CSP) as the number of unique values in the data is increased. Therefore, we suggest that our method may be a suitable alternative for experimenters looking to identify correlates of discrete independent variables with a large number of classes, as well as those of continuous independent variables.

Importantly, our method is also able to outperform canonical correlation analysis (CCA). CCA is very commonly used to identify

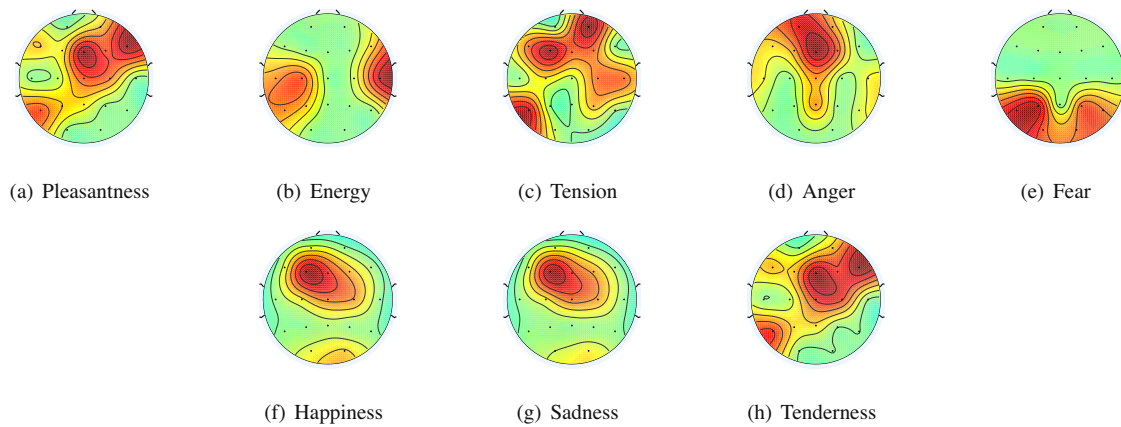


Fig. 5. Neural correlates of music induced affective states in the 10 Hz frequency band automatically identified by our method.

correlates of continuous variables and the superior performance of our method is, therefore, encouraging.

The classification accuracy achieved by our proposed method and CCA both exhibit higher variance in Fig. 1 than the CSP method. This could be due to the selection of a variable number of features by our proposed method. Specifically, CSP identifies a fixed number of filters, while the other methods select a number of features as a function of the training data.

There are a few potential reasons for the observed better performance of our proposed method to either CSP or CCA. First, CCA attempts to identify features which correlate with the independent covariate, whereas our method attempts to identify features which exhibit maximum variance with respect to the independent covariate. As a high correlation does not always equal high inter-group differences simply seeking features which correlate may not identify the best feature set. The variables our method identifies are more likely to exhibit larger inter-group differences.

Additionally, CSP is able to separate two groups in a dataset. However, for larger number of groups the performance declines. This is due to the decreasing numbers of items in each group leading to decreasing accuracy of the estimated covariance matrix.

Our method has been applied to identify neural correlates of music perception and music-induced affective states. The identified neural correlates of musical tempo were observed to lie over the left motor cortex. This is very similar to the manually identified neural correlates of musical tempo previously reported in Daly et al. (2014). The small difference in spatial location may be due to the small difference in frequency window size used between the studies.

The automatically identified neural correlates of music-induced affective states reveal relationships between EEG band-powers in the alpha band and each of the induced affective states. The prefrontal cortex is observed to be involved most frequently in music-induced affective state responses, and this is consistent with other findings (for example, Schmidt and Trainor, 2001; Schmidt and Hanslmayr, 2009; Dennis and Solomon, 2010). Additionally, affective states which are likely to be highly correlated with one another (for example, happiness and sadness) were observed to correlate with very similar spatial distributions of neural activity, lending further evidence to the efficacy and robustness of our method.

The neural correlates of anger and fear are interesting. These correlates are almost inverse maps of one another. Anger involves frontal cortical regions and central motor regions, while fear involves occipital regions and left/right motor regions. This observation warrants further investigation in future work.

For this study we only considered the 10 Hz frequency band of the EEG. The reason for this was due to the widely reported involvement of this frequency band in emotional responses (Kabuto et al., 1993; Schmidt and Trainor, 2001; Schaaff, 2008). Other frequency bands of the EEG have also been reported to be involved in emotional responses (Aftanas et al., 2006). However, as the primary purpose of our work is to demonstrate the efficacy of our proposed method, rather than to explore different neural correlates of emotion, we have restricted our search to the alpha frequency band.

Potential uses of our method are not limited to EEG studies. The method operates in feature space, hence any data set for which a correlation is sought between components of the dataset and a continuous or discrete independent variable containing multiple unique values may be explored via the use of this method. Additionally, the method could potentially be applied to identify features which differentiate groups of participants, for example different developmental groups.

However, it is important to consider a potential caveat of our method. The runtime and memory required by the method are proportional to the square of the size of the input data due to the use of covariance matrices due to the need to identify a similarity matrix and perform eigen-decomposition. Thus, for very large datasets it may be necessary to use parallelisation techniques to compute the results. However, this is also the case for CSP, which is also based upon eigen-decomposition.

In conclusion, the proposed method is able to perform an automated search for a subset of features in a dataset which relate to either discrete or continuous independent variables. It is able to outperform both CSP and CCA in terms of correlation between the independent covariate and the selected features as the number of groups increases and has a large number of potential applications. However, its use is currently restricted by the size of the dataset, although this restriction may be ameliorated by the use of advanced computation techniques such as parallel computing.

Acknowledgments

This work was supported by the EPSRC grants EP/J003077/1 and EP/J002135/1.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jneumeth.2014.12.012>.

References

- Aftanas LI, Reva NV, Savotina LN, Makhnev VP. Neurophysiological correlates of induced discrete emotions in humans: an individually oriented analysis. *Neurosci Behav Physiol* 2006;36(2):119–30, <http://dx.doi.org/10.1007/s11055-005-0170-6> <http://www.ncbi.nlm.nih.gov/pubmed/16380825>
- Allefeld C, Müller M, Kurths J. Eigenvalue decomposition as a generalized synchronization cluster analysis. *Int J Bifurc Chaos* 2007;17(10):3493–7, <http://dx.doi.org/10.1142/S0218127407019251>.
- Alpaydin E. Introduction to machine learning. Cambridge, MA: MIT Press; 2004 <http://books.google.com/books?id=1k0-WroiQEC&pgis=1>
- Daly I, Malik A, Hwang F, Roesch E, Weaver J, Kirke A, et al. Neural correlates of emotional responses to music: an EEG study. *Neurosci Lett* 2014;573:52–7, <http://dx.doi.org/10.1016/j.neulet.2014.05.003>.
- Daly I, Hallowell J, Hwang F, Kirke A, Malik A, Roesch E, et al. Changes in music tempo entrain movement related brain activity. In: *Proceedings of the EMBC*; 2014.
- David O, Friston K. A neural mass model for MEG/EEG: coupling and neuronal dynamics. *NeuroImage* 2003;20:1743–55.
- Dennis TA, Solomon B. Frontal EEG and emotion regulation: electro- cortical activity in response to emotional film clips is associated with reduced mood induction and attention interference effects. *Biol Psychol* 2010;85(3):456–64, <http://dx.doi.org/10.1016/j.biopsycho.2010.09.008>.
- Ellis DPW. Beat tracking by dynamic programming. *J New Music Res* 2007;36(1):51–60, <http://dx.doi.org/10.1080/09298210701653344>.
- Friedrich EV, Scherer R, Neuper C. The effect of distinct mental strategies on classification performance for brain–computer interfaces. *Int J Psychophysiol* 2012;84(1):86–94.
- Grosse-Wentrup M, Buss M. Multiclass common spatial patterns and information theoretic feature extraction. *IEEE Trans Bio-Med Eng* 2008;55(8):1991–2000, <http://dx.doi.org/10.1109/TBME.2008.921154> <http://www.ncbi.nlm.nih.gov/pubmed/18632362>
- Handy TC. Event-related potentials: a methods handbook. Cambridge, MA: MIT Press; 2005 <http://books.google.com/books?hl=en&lr=&id=OQyZEfgEzRUC&pgis=1>
- Hardle W, Simar L. Applied multivariate statistical analysis. Berlin, Heidelberg: Springer Berlin Heidelberg; 2007, <http://dx.doi.org/10.1007/978-3-540-72244-1> <http://www.springerlink.com/index/>
- Hardoon DR, Szedmak S, Shawe-Taylor J. Canonical correlation analysis: an overview with application to learning methods. *Neural Comput* 2004;16(12):2639–64, <http://dx.doi.org/10.1162/0899766042321814> <http://www.ncbi.nlm.nih.gov/pubmed/15516276>
- Hunter DB, Regan T. A note on the evaluation of the complementary error function. *Math Comput* 1972;26(118):539, <http://dx.doi.org/10.1090/S0025-5718-1972-0303685-4>.
- Hwang H-J, Kim S, Choi S, Im C-H. EEG-based brain–computer interfaces (BCIs): a thorough literature survey. *Int J Hum-Comput Int* 2013;29(12), <http://dx.doi.org/10.1080/10447318.2013.780869>, 130429122442009.
- Kabuto M, Kageyama T, Nitta H. EEG power spectrum changes due to listening to pleasant music and their relation to relaxation effects. *Nihon eiseigaku zasshi Jpn J Hyg* 1993;48(4):807–18 <http://www.ncbi.nlm.nih.gov/pubmed/8254987>
- Knapp TR. Canonical correlation analysis: a general parametric significance-testing system. *Psychol Bull* 1978;85(2):410–6.
- Koles ZJ, Lazar MS, Zhou SZ. Spatial patterns underlying population differences in the background EEG. *Brain Topogr* 1990;2(4):275–84, <http://dx.doi.org/10.1007/BF01129656> <http://link.springer.com/>
- Niedermeyer E, Silva FHL. *Electroencephalography: basic principles, clinical applications, and related fields*. Philadelphia, PA: Lippincott Williams & Wilkins; 2005.
- Pfurtscheller G, Lopes da Silva F. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol* 1999;110(11):1842–57, [http://dx.doi.org/10.1016/S1388-2457\(99\)00141-8](http://dx.doi.org/10.1016/S1388-2457(99)00141-8).
- Rahman M, Ma W, Tran D, Campbell J. A comprehensive survey of the feature extraction methods in the EEG research. In: Xiang Y, Stojmenovic I, Aduhan BO, Wang G, Nakano K, Zomaya A, editors. Algorithms and architectures for parallel processing, vol. 7440 of Lecture notes in computer science. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. p. 274–83, <http://dx.doi.org/10.1007/978-3-642-33065-0> <http://www.springerlink.com/index/>
- Schaaff K. EEG-based emotion recognition. Universität Karlsruhe; 2008 <http://www.citeulike.org/user/maxweizhao/article/6542875>
- Schmidt B, Hanslmayr S. Resting frontal EEG alpha-asymmetry predicts the evaluation of affective musical stimuli. *Neurosci Lett* 2009;460(3):237–40, <http://dx.doi.org/10.1016/j.neulet.2009.05.068>.
- Schmidt LA, Trainor LJ. Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cogn Emot* 2001;15(4):487–500, <http://dx.doi.org/10.1080/02699930126048>.
- Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. *Clin Neurophysiol* 2002;113:767–91.