Spatio-temporal analysis of feedback-related brain activity in brain-computer interfaces

Mahta Mousavi
Electrical and Computer Engineering
UC San Diego
La Jolla, CA, USA
mahta@ucsd.edu

Virginia R. de Sa
Department of Cognitive Science
UC San Diego
La Jolla, CA, USA
desa@ucsd.edu

Abstract

Electroencephalography-based brain-computer interfaces are systems that infer brain signals recorded using electroencephalography (EEG) to provide a means of communication for patients suffering from locked-in syndrome where common neuromuscular pathways are not available. One challenge in EEG-based BCIs is non-stationarity of the EEG signal. A major contributor to this is feedback-related brain activity. Since EEG consists of time series data recorded at multiple sites on the scalp, one can estimate covariance matrices for both time and space which lie on the Riemannian manifold of symmetric positive definite matrices. In this work, we investigate spatio-temporal aspects of the feedback-related brain activity by considering both space and time covariances in Euclidean space and on the Riemannian manifold. We propose two novel methods to incorporate both spatial and temporal features and show improved results compared to existing methods.

1 Introduction

Brain-computer interface (BCI) systems provide a means of communication for patients suffering from locked-in syndrome through bypassing the normal neuromuscular pathways [8, 18]. These systems record signals directly from the brain using methods such as electroencephalography (EEG) and infer the intention of the user from them. Motor imagery (MI) BCIs are common BCI paradigms in which the user imagines moving a part of her/his body. Movement imagination results in an event-related desynchronization (ERD) (decrease in power) in various frequency bands [16, 13]. Motor imagery of different body parts leads to different spatial desynchronization and this information is used by the BCI to distinguish among the imagined movement classes. In practice, for example, a user imagining moving her/his right or left hand can be mapped to a switch to turn on/off the light or to control the movement of a wheelchair, etc.

Non-stationarity of EEG signals is a barrier for real-world application of EEG-based BCI systems. Feedback-related brain activity is one contributor to this non-stationarity [6, 17, 14]. In this work, we investigate the spatio-temporal aspects of the feedback-related brain activity in a motor imagery BCI paradigm. Covariance matrices play an important role in extracting features for classification in MI-BCI [5, 10]. By investigating both space and time covariances in the Euclidean space and on the Riemannian manifold, our goal is to distinguish whether the BCI has made an error perceived by the user in a motor imagery BCI paradigm. This information can be used to improve the BCI performance (e.g. [14]).

Workshop on Modeling and Decision-Making in the Spatiotemporal Domain, 32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada.

2 Methods

Data were recorded from 10 participants participating in a motor imagery task to control a cursor on a monitor in front of them. Data were recorded using a 64-channel BrainAmp system (Brain Products GmbH). Each trial began with the cursor at the center of the screen and the target at either left or right side. The distance from the center to each side was fixed to three cursor jumps. The cursor moved at one step per second. The trial ended when the cursor hit the target or the other end of the screen. Participants were lead to believe they were in control of the cursor, but in reality, the cursor moved based on a pre-determined sequence that was the same for all participants. This was to have enough steps towards or away from the target irrespective of the participant's motor imagery performance. Data were downsampled and sections that were contaminated with excessive noise were removed. Next, independent component analysis (ICA) [11] was applied to remove muscle and eye artifacts. Pre-processing was done in MATLAB [12] and EEGLAB [7]. Feature extraction and classification were implemented in Python. We used the pyRiemann toolbox [2] to calculate Riemannian distances and means. For details on pre-processing, please refer to [14].

We looked at different frequency bands, namely 1-3, 2-5, 4-7, 6-10, 7-12, 10-15, 12-19, 18-25, 19-30, 25-35, 30-40 Hz to cover the low and high theta, mu, and beta frequency bands and to cover for potential individual differences [15]. Data were downsampled to 100 Hz and epoched at 100 to 1000 ms after each cursor movement which we call a "step". We trained a classifier to distinguish whether the user is satisfied with the last cursor movement or not, i.e., if the cursor had moved towards (good) or away (bad) from the target on that step. This is called a good/bad classifier. The methods that are proposed in this work have space and time covariances of these steps at their core.

2.1 Space and time covariances

Let $X_i \in \mathbb{R}^{c\times t}$ be an EEG epoch (step) where c is the number of channels, t represents the number of time samples and $i \in \{1, ..., N\}$ where N is the number of steps available. The sample space and time covariances for each step are defined as follows:

$$C_{s_i} = X_i X_i^T \qquad C_{t_i} = X_i^T X_i \tag{1}$$

Since the number of channels in our case $c = 64$ is smaller than the number of time samples (i.e., 90 time samples at a rate of 100 Hz), covariance matrices are not full rank and thus not positive definite. This is specifically required for the Riemannian methods. Therefore, we regularized both space and time covariances as follows:

$$C \leftarrow (1-\alpha)C + \alpha\frac{trace(C)}{N}I \tag{2}$$

where I is the identity matrix with the same size as C, α is the regularization parameter and $trace(C)$ is the sum of values on the diagonal of matrix C. In what follows, the regularization parameter was chosen arbitrarily but small, i.e., $\alpha = 0.001$. We did not test for other values. It is a part of our future work to use data-driven methods such as Ledoit-Wolf [9] to estimate the regularization parameter.

In this work, six methods are compared that differ in how they estimate the covariance from EEG data (either time or space) and how they extract corresponding features and infer the class label from this information. These methods are explained next.

2.2 Filterbank-Common Spatial Patterns (FB-CSP)

Filterbank common spatial patterns (FB-CSP) was proposed by by Ang et al. [1] and serves as our benchmark. In this method, first the average of the sample covariances for each of the good and bad classes in the Euclidean space were estimated as $\Sigma_g = \frac{1}{N}\sum_i C_{s_i}^g$ and $\Sigma_b = \frac{1}{N}\sum_i C_{s_i}^b$ respectively where N is the number of steps in the (balanced in number) good and bad classes. $C_{s_i}^g$ and $C_{s_i}^b$ represent the space covariance of the i^{th} good and bad steps respectively. CSP filters were estimated for each frequency band by simultaneous diagonalization of the two covariance matrices:

$$W^T\Sigma_g W = \Lambda^g \qquad W^T\Sigma_b W = \Lambda^b \tag{3}$$

where Λ^g and Λ^b are diagonal matrices and W is selected such that $\Lambda^g + \Lambda^b = I$ [5]. This can be achieved by solving the generalized eigenvalue problem: $\Sigma_g w = \lambda\Sigma_b w$. CSP filters $w_j, j = 1, ..., C$

are the columns of W. We considered the top 3 filters for each of the good and bad classes in each frequency band, i.e. the eigenvectors associated with the largest and smallest eigenvalues λ. Features were selected as the log of the variance of the EEG signal passed through these filters. A linear discriminant analysis (LDA) classifier was trained on the selected features [10].

2.3 Filterbank-Common Temporal Patterns (FB-CTP)

Common temporal patterns (CTP) were proposed by Yu et al. [19]. CTPs were found by solving the equations that were described earlier for CSP, except the sample mean of the good and bad time covariances were considered instead of space covariances. In their work, Yu et al. combined the features from common spatial and temporal filters. However, here we only considered the common temporal filters as a stand-alone method to compare with FB-CSP and another time covariance method on the Riemannian manifold that will be described next.

We consider a filter-bank version of CTP method similar to FB-CSP. The top 3 filters for each of the good and bad classes in each frequency band were selected. Features were the log of the variance of the EEG signal passed through these filters. A linear discriminant analysis (LDA) classifier was trained on the selected features [10].

2.4 FB-CSP-CTP

This method combines the spatial and temporal features from CSP and CTP methods. First CSP filters were trained and the top 3 filters for good and bad classes were selected. Then the EEG frames were passed through the 6 selected CSP filters. Next, CTP filters were trained and again the top 3 filters for each class were selected. The CSP filtered EEG frames passed through the selected CTP filters were used as features (6×6). This procedure was applied to each frequency band separately to have $11 \times 6 \times 6$ features and a logistic regression was trained on them for classification.

2.5 Filterbank-distance to spatial Riemannian means (FB-DRM-S)

Since covariance matrices are symmetric positive definite (SPD) matrices, they lie on a Riemannian manifold [3]. Let $P(n) = \{P \in S(n), P > 0\}$ be the set of all $n \times n$ SPD matrices. The Riemannian distance between two SPD matrices, P_1 and P_2 is defined as follows:

$$\delta_R(P_1, P_2) = ||Log(P_1^{-1}P_2)||_F = [\sum_{i=1}^{N} log^2 \lambda_i]^{1/2} \tag{4}$$

where λ_i are the eigenvalues of $(P_1^{-1}P_2)$ which are real positive (non-zero) values as P_1 and P_2 are both SPD. Also, $||.||_F$ is the Frobenius norm of a matrix and $Log(P)$ is the matrix logarithm of P. Since SPD matrices are diagonalizable, let $P = WDW^{-1}$. Then $Log(P) = Wlog(D)W^{-1}$ where $log(D)_i = log(D_i)$ and $log(D)_i$ and D_i are the i^{th} element on the diagonal of matrices $log(D)$ and D respectively.

The mean of the SPD matrices $P_1, P_2, ..., P_m$ on the Riemannian manifold is defined as follows [3]:

$$M(P_1, P_2, ..., P_m) = argmin_{P \in P(n)} \sum_{i=1}^{m} \delta_R^2(P, P_i). \tag{5}$$

Based on these mathematical tools on the Riemannian manifold, a filter bank generalization of the minimum distance to Riemannian mean (MDM) classifier [4] is proposed. First, the Riemannian mean of the good and bad space covariances on the training set were estimated. Then, the features were selected as the distances to the Riemannian means of good and bad classes in each frequency band (a total of 22 features: 11 frequency bands $\times$ 2 good and bad classes) and logistic regression was trained on the selected features.

2.6 Filterbank-distance to temporal Riemannian means (FB-DRM-T)

Here, we first estimated the Riemannian mean of the good and bad time covariances on the training set. Then, the features were selected as the distances to the Riemannian means of good and bad classes in each frequency band (a total of 22 features: 11 frequency bands $\times$ 2 good and bad classes) and logistic regression was trained on the selected features.

Table 1: G/B classification results.

ID	CSP	CTP	CSPCTP	DRM-S	DRM-T	DRM-ST
P1	0.73/0.017	0.65/0.014	0.81/0.011	0.72/0.012	0.68/0.016	0.74/0.019
P2	0.73/0.023	0.62/0.018	0.80/0.016	0.68/0.021	0.63/0.019	0.70/0.018
P3	0.57/0.020	0.54/0.017	0.59/0.020	0.59/0.021	0.58/0.015	0.58/0.019
P4	0.79/0.017	0.65/0.027	0.82/0.013	0.63/0.016	0.59/0.016	0.69/0.018
P5	0.62/0.014	0.55/0.013	0.75/0.016	0.69/0.019	0.62/0.019	0.71/0.020
P6	0.69/0.018	0.65/0.018	0.72/0.022	0.71/0.013	0.70/0.017	0.75/0.012
P7	0.73/0.014	0.60/0.016	0.70/0.017	0.69/0.023	0.67/0.024	0.73/0.019
P8	0.67/0.022	0.59/0.026	0.70/0.020	0.74/0.018	0.70/0.023	0.72/0.024
P9	0.75/0.020	0.67/0.021	0.66/0.014	0.82/0.018	0.76/0.021	0.81/0.018
P10	0.69/0.021	0.64/0.017	0.75/0.015	0.74/0.022	0.72/0.02	0.73/0.027
Average	0.70/0.021	0.62/0.014	0.73/0.023	0.70/0.020	0.66/0.019	0.72/0.018

2.7 Filterbank-distance to spatial and temporal Riemannian means (FB-DRM-TS)

This method combines the two previous ones: we first estimated the Riemannian mean of the good and bad time and space covariances on the training set. Then, the features were selected as the distances to the Riemannian means of good and bad time and space covariances in each frequency band (a total of 44 features: 11 frequency bands $\times$ 2 good and bad classes $\times$ 2 time and space covariances). A logistic regression classifier was trained on the selected features.

3 Results and Discussion

Since good and bad classes were not balanced, we sub-sampled the larger class to have balanced classes. We randomly sub-sampled the larger class and made 10 instances of train-validation and test sets - each having the same number of good and bad steps. Results are presented in table 1. The first column represents the participant ID (P1 to P10). In the rest of the columns, the first row specifies the classification method. Each entry shows the average classification accuracy on the test set across instances (first number) and standard error of the mean (second number). The last row shows the average performance of each method across participants. Our results show that FB-CSP-CTP performs best for majority of the participants. CSP is a very effective well developed way to remove the high amount of spatial (between electrodes) correlation in EEG data and it appears to be a good pre-processing step for future temporal analysis. On the other hand, FB-DRM-T on average performs better than FB-CTP. This is an interesting observation and implies that further investigation may be fruitful to best determine how to integrate Riemannian methods in an end-to-end classifier. Importantly in both the Riemannian methods and the CSP/CTP methods combining spatial and temporal information improves performance.

4 Conclusion

In this work, we investigated spatio-temporal aspects of the feedback-related brain activity. Since EEG is a time series recorded at multiple sites (electrodes on the scalp), we proposed methods to combine aspects of space and time covariances. We showed that time covariance can provide useful features for better classification of feedback-related brain activity. We proposed a filter-bank version of the common temporal patterns (CTP) method. Moreover, we proposed to combine spatial and temporal features by applying CTP after CSP for improved classification rates. We also proposed a filter bank generalization of the MDM method that was originally proposed in [4], i.e. our FB-DRM-S and another one to incorporate Riemannian distance to time covariances of the good and bad classes as extra features for the classifier (FB-DRM-TS). Future work includes investigating more sophisticated classifiers as introduced in [4], to replace the logistic regression on distance to the mean of each class in FB-DRM-TS to further improve our results.

Acknowledgments

This work was supported by NSF IIS 1219200, IIS 1817226, SMA 1041755, and IIS 1528214, FISP G2171, G3155, NIH 5T32MH020002-18 and Mary Anne Fox dissertation year fellowship.

References

[1] Kai Keng Ang, Zheng Yang Chin, Chuanchu Wang, Cuntai Guan, and Haihong Zhang. Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b. *Frontiers in neuroscience*, 6:39, 2012.

[2] Alexandre Barachant. pyriemann. `https://github.com/alexandrebarachant/pyRiemann`.

[3] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Riemannian geometry applied to bci classification. In *International Conference on Latent Variable Analysis and Signal Separation*, pages 629–636. Springer, 2010.

[4] Alexandre Barachant, Stéphane Bonnet, Marco Congedo, and Christian Jutten. Multiclass brain–computer interface classification by riemannian geometry. *IEEE Transactions on Biomedical Engineering*, 59(4):920–928, 2012.

[5] Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, and K-R Muller. Optimizing spatial filters for robust eeg single-trial analysis. *IEEE Signal processing magazine*, 25(1):41–56, 2008.

[6] Ricardo Chavarriaga, Aleksander Sobolewski, and José del R Millán. Errare machinale est: the use of error-related potentials in brain-machine interfaces. *Frontiers in neuroscience*, 8:208, 2014.

[7] Arnaud Delorme and Scott Makeig. EEGlab: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neuroscience Methods*, 134(1):9–21, 2004.

[8] Lawrence Ashley Farwell and Emanuel Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70(6):510–523, 1988.

[9] Olivier Ledoit and Michael Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5):603–621, 2003.

[10] Fabien Lotte, Laurent Bougrain, Andrzej Cichocki, Maureen Clerc, Marco Congedo, Alain Rakotomamonjy, and Florian Yger. A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update. *Journal of neural engineering*, 15(3):031005, 2018.

[11] Scott Makeig, Anthony J Bell, Tzyy-Ping Jung, and Terrence J Sejnowski. Independent component analysis of electroencephalographic data. In *Advances in neural information processing systems*, pages 145–151, 1996.

[12] MATLAB and Statistics Toolbox Release 2015a. The MathWorks Inc., Natick, Massachusetts, United States, 2015.

[13] Dennis J McFarland, Laurie A Miner, Theresa M Vaughan, and Jonathan R Wolpaw. Mu and beta rhythm topographies during motor imagery and actual movements. *Brain topography*, 12(3):177–186, 2000.

[14] Mahta Mousavi, Adam S Koerner, Qiong Zhang, Eunho Noh, and Virginia R de Sa. Improving motor imagery bci with user response to feedback. *Brain-Computer Interfaces*, 4(1-2):74–86, 2017.

[15] Eunho Noh, Grit Herzmann, Tim Curran, and Virginia R de Sa. Using single-trial eeg to predict and analyze subsequent memory. *NeuroImage*, 84:712–723, 2014.

[16] Gert Pfurtscheller, Clemens Brunner, Alois Schlögl, and FH Lopes Da Silva. Mu rhythm (de) synchronization and eeg single-trial classification of different motor imagery tasks. *NeuroImage*, 31(1):153–159, 2006.

[17] Gerwin Schalk, Jonathan R Wolpaw, Dennis J McFarland, and Gert Pfurtscheller. Eeg-based communication: presence of an error potential. *Clinical neurophysiology*, 111(12):2138–2144, 2000.

[18] Jonathan R Wolpaw, Dennis J McFarland, Gregory W Neat, and Catherine A Forneris. An eeg-based brain-computer interface for cursor control. *Electroencephalography and clinical neurophysiology*, 78(3):252–259, 1991.

[19] Ke Yu, Kaiquan Shen, Shiyun Shao, Wu Chun Ng, Kenneth Kwok, and Xiaoping Li. Common spatio-temporal pattern for single-trial detection of event-related potential in rapid serial visual presentation triage. *IEEE Transactions on Biomedical Engineering*, 58(9):2513–2520, 2011.