Improving Robustness in Motor Imagery Brain-Computer Interfaces

Mahta Mousavi
Electrical Engineering and Computer Science
TU Berlin, Germany
mahta@ucsd.edu

Eric Lybrand
Department of Mathematics
UC San Diego, CA, USA
elybrand@ucsd.edu

Shuangquan Feng
Department of Cognitive Science
UC San Diego, CA, USA
s1feng@ucsd.edu

Shuai Tang
AWS AI Labs
New York, NY, USA
shuat@amazon.com

Rayan Saab
Department of Mathematics
Halıcıoğlu Data Science Institute
UC San Diego, CA, USA
rsaab@ucsd.edu

Virginia R. de Sa
Department of Cognitive Science
Halıcıoğlu Data Science Institute
UC San Diego, CA, USA
desa@ucsd.edu

Abstract

The method of Common Spatial Patterns (CSP) is widely used for feature extraction of electroencephalography (EEG) data such as in motor imagery brain-computer interface (BCI) systems. It is a data-driven method estimating a set of spatial filters so that the power of the filtered EEG data is maximally separated between imagery classes. This method, however, is prone to overfitting and is known to suffer from poor generalization especially with limited calibration data. In this work, we propose a novel algorithm called *Spectrally Adaptive Common Spatial Patterns* (SACSP) that improves CSP by learning a temporal/spectral filter for each spatial filter so that the spatial filters are concentrated on the most relevant temporal frequencies for each user. We show the efficacy of SACSP in motor imagery BCI in providing better generalizability and higher classification accuracy from calibration to online control compared to existing methods while providing neurophysiologically relevant information about the temporal frequencies of the filtered signals.

Introduction:The method of Common Spatial Patterns (CSP) [1] is among the most commonly used methods for feature extraction in electroencephalography (EEG) data [2, 3]. One of its main applications is to learn relevant features in motor imagery brain-computer interface (MI-BCI) systems [4]. Imagined movement, similar to actual movement, results in a decrease in power in user specific frequency ranges within the mu (7-13 Hz) and often beta (13-30Hz) frequency bands over the contralateral sensorimotor cortex. CSP is a data-driven supervised approach that learns a set of spatial filters to maximally discriminate the two classes. That is, the power of the projected EEG data through the selected filters is maximally different for the two imagery classes [5]. CSP counters the spatial blurring in EEG data caused by measuring electrical activity through bone and several layers of tissue [5].

During the calibration session in MI-BCI, the user is instructed to imagine the movement of different body parts (e.g., the right and the left hand) and is provided with no feedback or sham feedback. The recorded EEG data during this calibration session – calibration data – is then used to train the classifier (CSP filters + linear classifier [5]) and deployed for real-time online control of the BCI. However, CSP suffers from poor generalizability from calibration to online control [6].

35th Conference on Neural Information Processing Systems (NeurIPS 2021), Sydney, Australia.

While CSP learns the user-specific spatial filters, it does not learn the spectral filters and equally weights the spectral information within the 7-30 Hz frequency band. The Canonical correlation approach to common spatial patterns (CCACSP) [7] combines canonical correlation analysis (CCA) and CSP approaches and its cost function considers a temporal filter with a forward and backward one sample temporal shift. This results in a weighted spectral filter that emphasizes the lower frequencies, but this method does not learn user specific frequency bands. Spectrally-weighted CSP (spec-CSP) attempts to find a set of user specific spatial and spectral (frequency transformed temporal) filters through alternating optimization of the cost function [8] but it is computationally expensive and does not result in robust filters in practice.

Instead of learning the optimal temporal/spectral filters, another line of CSP variants finds the relevant sub-bands through a filter bank approach [9, 10]. These methods require a priori assumptions on the length and location of the sub-bands and require significant amounts of calibration data to find a robust feature space to be transferred from calibration data to online BCI control [11].

Deep learning methods have been developed for EEG data as well. EEGNet is one such technique that is widely used for EEG data [12] that attempts at learning user specific temporal and spatial filters in an end-to-end fashion; however, the major limitation of deep neural networks for EEG-based BCI is that such networks have a very large number of parameters, which thus requires a very large number of training examples to calibrate them. In practice, the access to calibration data from each user is limited.

To overcome the limited generalizability of CSP and the approaches mentioned above, we propose a novel approach to learn spatial and spectral filters to emphasize the most relevant frequencies for each user in motor imagery data by solving an optimization problem that generalizes the cost function in CSP. We apply our method to motor imagery data that we recorded from 12 participants and show improved classification accuracy as well as more interpretable spatial patterns and spectral filters compared to the closely related methods of CSP [1], CCACSP [7], spec-CSP [8] and a deep learning method developed for EEG data called EEGNet [12].

Methods: CSP estimates the spatial covariance for the two classes and uses a Rayleigh quotient cost function to find the optimal directions for such projection [1, 5]. In SACSP however, we define the spectrally-weighted spatial covariance instead, to weight the motor imagery related frequencies.

Let $X_i \in \mathbb{R}^{n\times t}$ represent the bandpassed EEG data epoch over n channels with t time samples that belongs to class $i \in \{1, 2\}$, e.g. the two motor imagery classes. Since EEG data is bandpassed, $\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0$. Let C be a circulant matrix. We define the spectrally-weighted spatial covariance matrices for classes 1 and 2 as $\Gamma_{S1} = \mathbb{E}[X_1 C X_1^T]$ and $\Gamma_{S2} = \mathbb{E}[X_2 C X_2^T]$, where $(\cdot)^T$ indicates the transpose of a matrix/vector. As C is circulant, we can write it as $C = F\text{diag}(h)F^H$ where h is a vector, $diag(h)$ is a matrix whose diagonal elements are the elements of the vector h, F is the $t \times t$ Fourier matrix and $(\cdot)^H$ indicates the complex conjugate transpose of a matrix. Replacing this in the above and defining $\hat{X_1} = X_1 F$ and $\hat{X_2} = X_2 F$ as the Fourier transform of an epoch along its temporal dimension we can write the following set of optimization problems:

$$\underset{w,h}{\arg\max} \quad \frac{w^T \mathbb{E}[\hat{X}_1 \text{diag}(h) \hat{X}_1^H] w}{w^T (\Sigma_1 + \Sigma_2) w}, \qquad \underset{v,l}{\arg\max} \quad \frac{v^T \mathbb{E}[\hat{X}_2 \text{diag}(l) \hat{X}_2^H] v}{v^T (\Sigma_1 + \Sigma_2) v}, \tag{1}$$

where $\Sigma_i = \mathbb{E}[X_i X_i^T]$ is the spatial covariance of each class. We solve the two optimization problems in (1) via an alternating maximization procedure.

Classification: After finding the optimal $2 \times R$ (where R is the number of filters for each class) spatial and spectral filter pairs (f_{wv}, f_{hl}), then $2 \times R$ features were extracted by filtering each epoch, x as follows:

$$\log(f_{wv}^T \, (\hat{x} \, \text{diag}(f_{hl}) \hat{x}^H) \, f_{wv}), \tag{2}$$

where $\hat{x} = xF$. Next, linear discriminant analysis (LDA) [13] with automatic shrinkage using the Ledoit-Wolf lemma [14] was trained as the classifier.

Data: Participants were instructed to use movement imagination of their right or left hand to move a cursor on the screen. Each participant participated in 9 blocks with 20 trials in each block. Each trial began with a cursor (blue circle about 2 cm in diameter) in the center and a target (white circle about 2 cm in diameter) at either side of the screen and three steps away from the cursor. The cursor moved one step per 1.2 seconds towards or away from the target until it reached the target location or the corresponding location on the other side of the screen.

During the first three blocks, the cursor moved according to a pre-determined pseudo-randomly generated sequence of movements but the participants were not aware of this and were led to believe that they were in control of the cursor movements. The recorded EEG data from the first three blocks were used for calibration. CSP was trained for feature extraction where the top three CSP filters for each class were selected. The logarithm of the power of the CSP-filtered EEG data were selected as features and shrinkage-LDA [14] was trained for classification. This classifier was used during the online control of the cursor in the latter 6 blocks which had a similar structure as the calibration data except that the cursor moved based on the participant's motor imagery data.

In offline analysis, we used the data from the calibration blocks to train CSP, CCACSP, spec-CSP, EEGNet and SACSP separately. Each method was then tested on data recorded during online control blocks.

Pre-processing: Data from calibration and online blocks were downsampled to 100 Hz and epoched 0-1 second with respect to each cursor movement (step) excluding the last. After downsampling and epoching, data were re-referenced to the common average and bandpass filtered in 7–30 Hz to include the mu (about 7–13 Hz) and beta (about 13–30 Hz) frequency bands [15, 16].

We balanced the right and left classes (including movements towards and away from the target in each class) for both calibration and online blocks to alleviate any potential bias in learning features and for easier interpretation of the results. Balancing the classes was done 10 times to allow steps to be in the training (from calibration data) or test (from online data) sets.

Classification accuracy: The classification accuracy of the CSP, CCACSP, spec-CSP, EEGNet and SACSP when the feature extraction and classifiers were trained on the calibration data and tested on the online data, are presented in table 1. The first number in each entry for participants 1 to 12 (P1-P12), represents the mean classification accuracy over 10 instances of train-test data and the second number reports the standard deviation. In the last row, the average performance across participants is reported where in each entry, the first number indicates the average for the corresponding column and the second number indicates the standard error of the mean. We should emphasize that the classification accuracy is for only one second of motor imagery data (one cursor movement or step) and still more than half of the participants perform around 0.7 or higher with the proposed algorithm. Across participants, SACSP performs significantly better compared to CCACSP, spec-CSP, CSP and EEGNet (Wilcoxon signed rank test, $p < 0.01$).

We also looked at the classification accuracy when each method was trained on calibration data and tested on calibration data through cross-validation (we call it 'calibCV' case) as well as when trained on online data and tested on online data through cross-validation (we call it 'onlineCV' case). Note that both of these conditions are different from the practical scenario of interest in this work, where the methods are trained on the calibration data and tested on the online data (we call it 'test' case). Our results show that for CSP, calibCV and onlineCV are not different across participants (average 0.73 and 0.71, Wilcoxon signed-rank test, $p = 0.42$). However, there is a large drop in performance of CSP in the test condition (across participants significantly lower that calibCV and onlineCV, Wilcoxon signed rank test, $p < 0.001$). Our results show that spec-CSP, CCACSP, and SACSP were all able to close this gap from calibration to online performance with SACSP performing significantly better than all other methods (Wilcoxon signed rank test, $p < 0.01$).

Visualization: We visualize the optimal spatial patterns and spectral filters to compare their neurophysiological interpretation. Normalized spectral filters were plotted against frequency (Hz) as line plots using MATLAB [17]. Spatial patterns were estimated for each method accordingly and `topoplot` from EEGLAB [18] was used to plot them as heatmaps on a 2-D scalp. Figure 2 compares the spatial patterns and spectral filters of the SACSP with CSP, CCACSP, and spec-CSP in one of the participants (P1). Note that these are the filters and patterns that were trained on the calibration data and used later for classification of the online data. The last two rows show the results for SACSP. The filters and patterns for each class (Pw_1, Pw_2 and Pw_3, h_1, h_2 and h_3) and (Pv_1, Pv_2 and Pv_3, l_1, l_2 and l_3) converge to similar values. This is because our algorithm, different from the others presented here, does not result in spatial filters/patterns that are necessarily orthogonal to each other and convergence to the same solution from different initializations in fact shows that SACSP converged to a robust solution.Furthermore, we tested SACSP with random initialization and received the same filter and patterns.

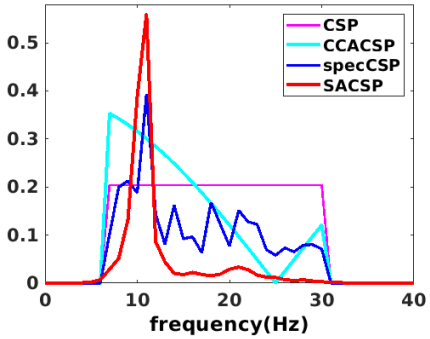


Figure 1: Comparison of the normalized spectral filters in one of the participants (P9). SACSP and spec-CSP lines show the top spectral filter for the right hand motor imagery class. Note that CSP and CCACSP have fixed spectral filters.

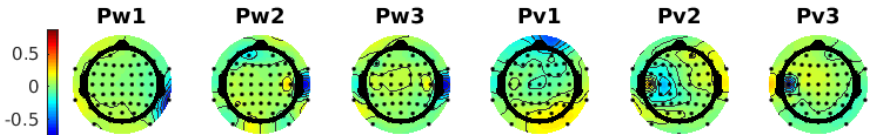
Table 1: Motor imagery (right/left hand) classification accuracy. For P1-P12, the first number in each entry reports the average accuracy across 10 instances of the data and the second number reports the standard deviation of the mean. In the last row, the first number in each entry reports the average classification accuracy across participants for the corresponding column and the second number reports the standard error of the mean.

PID	CSP	spec-CSP	CCACSP	EEGNet	SACSP
P1	0.55/0.035	0.61/0.039	0.74/0.033	0.61/0.032	0.81/0.022
P2	0.56/0.028	0.63/0.091	0.69/0.012	0.56/0.051	0.69/0.030
P3	0.63/0.038	0.61/0.037	0.61/0.057	0.60/0.074	0.72/0.096
P4	0.72/0.019	0.70/0.051	0.71/0.020	0.60/0.028	0.73/0.009
P5	0.59/0.118	0.59/0.105	0.56/0.071	0.56/0.058	0.62/0.117
P6	0.50/0.014	0.52/0.032	0.54/0.017	0.52/0.016	0.52/0.013
P7	0.62/0.027	0.65/0.035	0.68/0.019	0.58/0.044	0.70/0.040
P8	0.57/0.023	0.59/0.034	0.71/0.023	0.56/0.050	0.72/0.036
P9	0.68/0.048	0.67/0.033	0.73/0.019	0.69/0.023	0.76/0.023
P10	0.49/0.021	0.49/0.024	0.49/0.017	0.52/0.027	0.50/0.019
P11	0.60/0.019	0.64/0.028	0.64/0.034	0.61/0.023	0.67/0.027
P12	0.47/0.032	0.45/0.025	0.49/0.024	0.49/0.026	0.51/0.010
Average	0.58/0.022	0.60/0.022	0.63/0.027	0.58/0.015	**0.66/0.030**

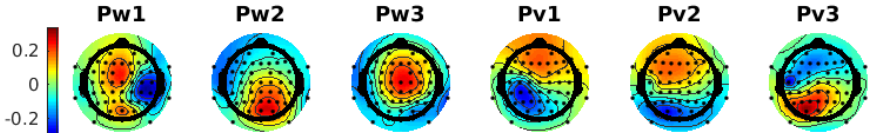
Comparing the spectral filters for spec-CSP and SACSP, we can see that the spec-CSP spectral filters fail at selectively extracting the underlying features in an interpretable way. SACSP shows smoother and less noisy filters with a peak at/around the mu band for both the right and left hand motor imagery classes and a small beta peak for the left hand motor imagery class. Another example comparing spectral filters for another participant (P9) is presented in figure 1.

SACSP shows spatial patterns typical for right and left hand motor imagery [4, 15]. Pv_3 in spec-CSP can be regarded as a similar pattern to Pv_3 in SACSP but the rest of the patterns in spec-CSP and those in CSP are neurophysiologically implausible for motor imagery. Pv_1 and Pv_3 in CCACSP are motor imagery patterns. However, CCACSP fails at finding appropriate patterns for the other class.

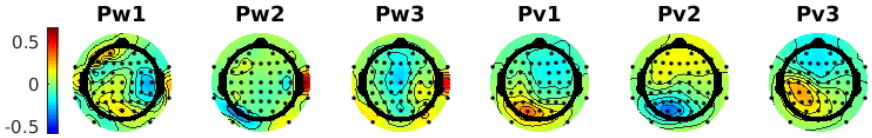
Discussion and Conclusion: Most work in the literature focuses on cross-validation accuracy (not the practical use case of calibration to online transfer) and over longer lengths of motor imagery data. In this work, we proposed SACSP and showed its superiority in classification accuracy when trained on limited amount of calibration data and tested on online data as well as neurophysiologically plausible spatial patterns and spectral filters.



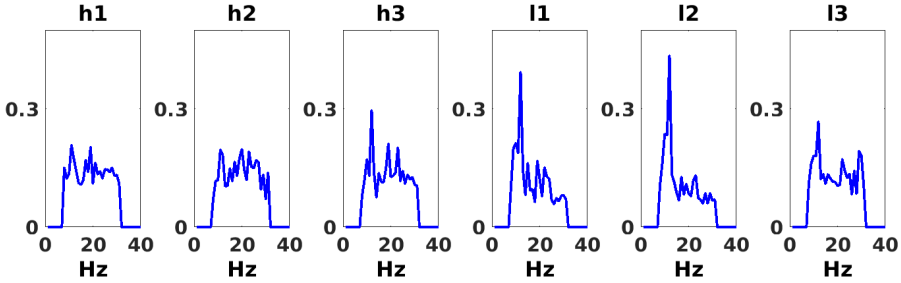
(a) P1 – CSP spatial patterns.



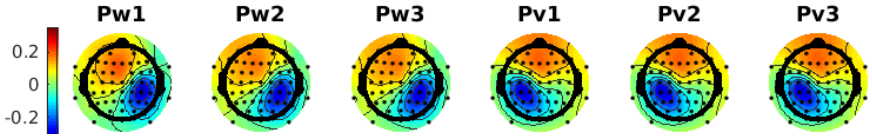
(b) P1 – CCACSP spatial patterns.



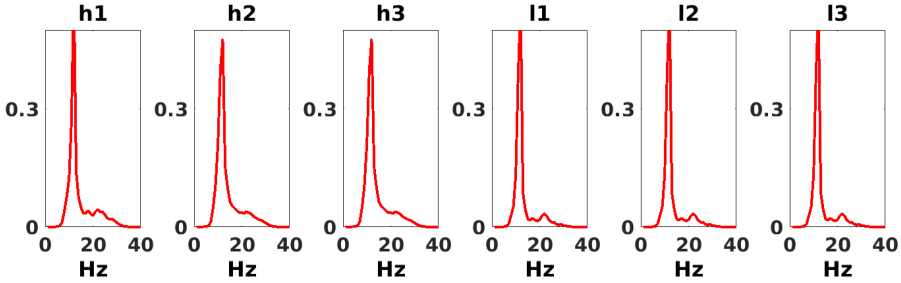
(c) P1 – spec-CSP spatial patterns.



(d) P1 – spec-CSP spectral filters.



(e) P1 – SACSP spatial patterns.



(f) P1 – SACSP spectral filters.

Figure 2: Spatial patterns and spectral filters trained on the calibration data for one of the participants (P1). Compared to other methods, SACSP results in spatial patterns that are typical for right and left hand motor imagery while providing neurophysiologically relevant frequency selectivity in spectral filters. Normalized spectral filters were plotted against frequency (Hz) as line plots using MATLAB. Spatial patterns were estimated for each method accordingly and `topoplot` from EEGLAB was used to plot them as heatmap on a 2-D scalp.

References

[1] Herbert Ramoser, Johannes Muller-Gerking, and Gert Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering*, 8(4):441–446, 2000.

[2] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. Brain computer interfaces, a review. *Sensors*, 12(2):1211–1279, 2012.

[3] Piotr Wierzgała, Dariusz Zapała, Grzegorz M Wojcik, and Jolanta Masiak. Most popular signal processing methods in motor-imagery BCI: a review and meta-analysis. *Frontiers in neuroinformatics*, 12:78, 2018.

[4] Gert Pfurtscheller and Christa Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134, 2001.

[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] Pradeep Shenoy, Matthias Krauledat, Benjamin Blankertz, Rajesh PN Rao, and Klaus-Robert Müller. Towards adaptive classification for BCI. *Journal of Neural Engineering*, 3(1):R13, 2006.

[7] Eunho Noh and Virginia R De Sa. Canonical correlation approach to common spatial patterns. In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 669–672. IEEE, 2013.

[8] Ryota Tomioka, Guido Dornhege, Guido Nolte, Benjamin Blankertz, Kazuyuki Aihara, and Klaus-Robert Müller. Spectrally weighted common spatial pattern algorithm for single trial EEG classification. *Dept. Math. Eng., Univ. Tokyo, Tokyo, Japan, Tech. Rep*, 40, 2006.

[9] Kai Keng Ang, Zheng Yang Chin, Haihong Zhang, and Cuntai Guan. Filter bank common spatial pattern (FBCSP) in brain-computer interface. In *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, pages 2390–2397. IEEE, 2008.

[10] 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.

[11] 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.

[12] Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of neural engineering*, 15(5):056013, 2018.

[13] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct):2825–2830, 2011.

[14] Olivier Ledoit and Michael Wolf. Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4):110–119, 2004.

[15] 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.

[16] 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.

[17] MATLAB and Statistics Toolbox Release 2018b. The MathWorks Inc., Natick, Massachusetts, United States, 2018.

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