Class-distinctiveness-based frequency band selection on the Riemannian manifold for oscillatory activity-based BCIs: preliminary results

Maria Sayu Yamamoto, Fabien Lotte, Florian Yger, Sylvain Chevallier

Class-distinctiveness-based frequency band selection on the Riemannian manifold for oscillatory activity-based BCIs: preliminary results

Maria Sayu Yamamoto[1], Fabien Lotte[2], Florian Yger[3] and Sylvain Chevallier[1,4]

***Abstract*— Considering user-specific settings is known to enhance Brain-Computer Interface (BCI) performances. In particular, the optimal frequency band for oscillatory activity classification is highly user-dependent and many frequency band selection methods have been developed in the past two decades. However, it is not well studied whether those conventional methods can be efficiently applied to the Riemannian BCIs, a recent family of BCI systems that utilize the non-Euclidean nature of the data unlike conventional BCI pipelines. In this paper, we proposed a novel frequency band selection method working on the Riemannian manifold. The frequency band is selected considering the class distinctiveness as quantified based on the inter-class distance and the intra-class variance on the manifold. An advantage of this method is that the frequency bandwidth can be adjusted for each individual without intensive optimization steps. In a comparative experiment using a public dataset of motor imagery-based BCI, our method showed a substantial improvement in average accuracy over both a fixed broad frequency band and a popular conventional frequency band selection method. In particular, our method substantially improved performances for subjects with initially low accuracies. This preliminary result suggests the importance of developing new user-specific setting algorithms considering the manifold properties, rather than directly applying methods developed prior to the rise of the Riemannian BCIs.**

I. INTRODUCTION

An oscillatory activity-based Brain-Computer Interface (BCI) is a system that utilizes oscillations in brain signals to operate external devices [1]. Oscillatory activity can be observed in various brain areas and frequency bands. Among them, the sensorimotor rhythms (SMRs) are the ones that are recorded over the sensori-motor cortex in the α and β bands [2]. SMRs can be modulated by execution or imagination of motor activity, which can be used for motor imagery (movement imagination) BCIs (MI-BCIs). A practical example of an MI-BCI is a system to control a computer cursor by imagining hand movements [3]. The user voluntarily imagines hand movement without actually moving, which leads to distinguishable changes in SMRs over the contralateral motor and sensori-motor cortices. This activity is typically measured by electroencephalography (EEG) and translated into the corresponding BCI command, *i.e.,* move a cursor towards left/right on a screen, if the user performs left/right-hand MI. Therefore, there is a need for BCIs that can accurately recognize which MI tasks the user performed, to obtain reliable BCI applications.

In this context, it is known that BCI performance can be enhanced by considering user-specific settings [4]. In particular, the choice of the frequency band to use is an important parameter. SMRs modulations are typically observed in the α and β bands [2], however, the most modulated frequency band is different for each individual [5]. Within the past two decades, a number of strategies have been proposed for user-specific frequency band selection, such as heuristic-based methods [6], filter-bank methods [7] and probabilistic methods [8]. Recently, it was also reported that there was a correlation between MI-BCI performance and user-specific frequency band characteristics [9].

Meanwhile, the use of Riemannian geometry has also contributed to improve BCI reliability [10], [11], [12]. Instead of conventional processing pipelines such as estimating spatial filters and selecting features, the distinguishing point of Riemannian BCIs is describing the EEG data by sample covariance matrices (SCMs) and utilizing geometrical properties of the space where they belong, called the Riemannian manifold, for classification. Riemannian BCIs showed superior to many other conventional processing pipelines in many datasets and competitions [13], [11].

User specific settings should be beneficial to Riemannian BCIs as well. However, it has not been studied systematically whether conventional frequency band selection methods can provide optimal bands for Riemannian classifiers as well. While a method has been proposed to select among a few frequency bands with a fixed length for Riemannian classifiers [14], there is, to the best of our knowledge, no method to fully optimize (both in location and length) the frequency band selection for Riemannian BCIs.

In this paper, we propose a novel frequency band selection method that works on the Riemannian manifold. The frequency band is automatically selected by evaluating its class distinctiveness on the manifold. The class distinctiveness of each frequency band is quantified by the ratio of inter-class distance to intra-class variance, and the most distinct frequency bands are selected and concatenated using a heuristic to define an appropriate threshold for each individual. To the best of our knowledge, this is the first frequency band optimization method that works on the manifold of EEG covariance matrices. An advantage of our method is that the frequency bandwidth can be adjusted for each individual without any heavy optimization process; therefore, it should

*MS.Y was supported by the ANR UDOPIA grant. F.L. was supported by the ERC with project BrainConquest (grant ERC-2016-STG-714567).

[1]LISV, Université Paris-Saclay, 78140, Vélizy-Villacoublay, France maria-sayu.yamamoto@universite-paris-saclay.fr

[2]Inria Bordeaux Sud-Ouest / LaBRI (Univ. Bordeaux / CNRS / Bordeaux INP), 33405, Talence, France fabien.lotte@inria.fr

[3]LAMSADE, PSL-Université Paris-Dauphine, 75016, Paris, France florian.yger@dauphine.fr

[4]Université de Versailles Saint-Quentin-en-Yvelines, 78035, Versailles, France sylvain,chevallier@uvsq.fr

be practical for real BCI applications. In the experiment with a Riemannian classifier, our method showed a substantial improvement in average accuracy compared to a fixed-frequency band selection and a conventional frequency band selection method [6] that is not based on SCM features. This preliminary result suggests the effectiveness of frequency band selection considering the geometrical features on a manifold for Riemannian BCIs.

This paper is organized as follows: Section II describes the principles of Riemannian geometry and Section III describes our new approach. The experimental evaluation is described in section IV. Then, Section V describes the results and discussion, while Section VI concludes the paper.

II. Riemannian geometry

In this section, we briefly introduce Riemannian geometry from a practical point of view.

Let $X \in \mathbb{R}^{M \times T}$ be a filtered EEG trial, with M channels and T time samples. The SCM of X is defined as $C_X = \frac{1}{T-1} X X^\top$. The diagonal entries of this matrix represent the variance of each channel signal, *i.e.,* the band power of each channel, while the off-diagonal terms contain the covariance between each pair of channels. This estimated covariance matrix is empirically symmetric positive definite (SPD), *i.e.,* it can be always diagonalized with strictly positive eigenvalues. The set of SPD matrices can be analyzed with a particular geometry making it the Riemannian manifold. On such curved spaces, distance between matrices are computed along geodesic. The affine-invariant Riemannian metric (AIRM) [15] distance is a geodesic measurement that respects the original curvature of the manifold and it is given as follows:

$$\delta_r(C_1, C_2) = \|\log(C_1{}^{-\frac{1}{2}} C_2 C_1{}^{-\frac{1}{2}})\|_F = \left(\sum_{i=1}^{n} \log^2 \lambda_i \right)^{1/2} \quad (1)$$

where λ_i are the positive eigenvalues of $C_1{}^{-\frac{1}{2}} C_2 C_1{}^{-\frac{1}{2}}$ and $\|.\|_F$ is the Frobenius norm of the matrix.

A basic and well-known Riemannian classifier is the Minimum Distance to the Riemannian Mean (MDRM) [10]. With this classifier, a new observation is classified based on the nearest class centroid with the shortest AIRM distance. Each class centroid is estimated by the Riemannian mean of a set of training SPD matrices, see [10, Algorithm.1].

III. Proposed method

Our frequency band selection method combines a filter bank with a heuristic based on class separability on a Riemannian manifold. Here, we describe how to measure class separability on a manifold and our algorithm.

A. Class Distinctiveness

To measure how distinct the SCM patterns are between classes on a manifold, we used a metric called *classDis* [16]. Let $C^K = [C_1, C_2, \cdots, C_n]$ be a set of SCMs of class K ($K = A, B$). The classDis between two classes A and B for a set of SCMs filtered in a certain frequency band f is given as:

$$\text{classDis}(f) = \frac{\delta_r\left(\overline{C^A}, \overline{C^B}\right)}{\frac{1}{2}\left(\sigma_{C^A} + \sigma_{C^B}\right)} \quad (2)$$

where $\overline{C}$ is the Riemannian mean and σ_C is the Riemannian mean absolute deviation given by $\sigma_C = \frac{1}{n}\sum_{i=1}^{n} \delta_r\left(C_i, \bar{C}\right)$ for the corresponding class. The numerator of this equation indicates how far apart the mean SCMs of each class are. The denominator represents the degree of dispersion around the mean (*i.e.,* variance) within each class. Thus, high classDis value are obtained with a large distance between class means and a small variance within class.

B. Algorithm

We describe the whole procedure of our method in pseudo-code, see Algorithm 1.

First, training raw EEG data are filtered for each sub-band of a filter bank. Here, we used a total of $S = 14$ sub-bands with 4 Hz bandwidth and 2 Hz overlap, *i.e.* 5–9, 7–11, $\cdots$, 31–35 Hz. Next, SCMs for each sub-band are estimated. Note that we used the regularized SCMs with the oracle approximating shrinkage [17]. Then, the class distinctiveness of each sub-band are evaluated by using the classDis metric. Using the sub-band with maximum classDis f^{max} as a starting point, the selected frequency band is then enlarged as long as the classDis value exceeds a threshold of $\max(\text{classDis}) - (\max(\text{classDis}) - \min(\text{classDis})) \times 0.4$, determined empirically. Since the goal is to select the frequency bands with relatively high classDis values for each subject, we assumed that it was reasonable to select values falling into the top 40% of the classDis range. Finally, all selected sub-bands are concatenated so that a range from the smallest to the largest frequency of the concatenated sub-bands $[f_0^{\text{start}}, f_1^{\text{end}}]$ is selected as the optimal user-specific frequency band.

Algorithm 1 Frequency band selection with classDis metric

Input: N training raw EEG data $X_{\text{train}} = [X_1, X_2, \cdots, X_N]$
Input: S different sub-band $f^{(s)} = [f_0{}^s, f_1{}^s] (s = 1, 2, \cdots, S)$
Output: Subject specific frequency band $[f_0, f_1]$

```
Filter X_train with S different sub-band
Estimate SCMs per sub-band C^(s) = [C_1^s, C_2^s, ..., C_N^s]
Compute classDis(f^(s)) per sub-band
Find the sub-band with maximum classDis value
    f^max <- argmax_s classDis(f^(s))
Initialize: f_0^start <- f_0^max    f_1^end <- f_1^max
Threshold: (max(classDis) - min(classDis)) × 0.4
while classDis(f^start - 1) >= threshold do
    f^start <- f^start - 1
end while
while classDis(f^end - 1) >= threshold do
    f^end <- f^end + 1
end while
    return f_0 = f_0^start,    f_1 = f_1^end
```

IV. Experimental evaluation

We evaluated our method by comparing classification accuracy with a baseline that does not apply any frequency band selection but uses a fixed neurophysiologically relevant band for all subjects, and with a conventional method that is a well-known selection method based on EEG spectral power. In this section, we summarize the experimental setups.

A. Data description

We used dataset IIa from BCI competition IV, provided by TU Graz, Austria [18]. This dataset comprises EEG signals from 9 subjects who performed left hand, right hand, both feet, and tongue MI. EEG were recorded using 22 channels. Training and testing sets are available for each subject. Both sets contain 72 trials lasting 7.0 s for each class. In this work, we only used right and left hand MI EEG data. From training data, one epoch was extracted from 0.5 to 2.5 s after the instruction cue (at $t = 2.0$ s) for each trial. In the testing data, sliding windows were used to simulate an online classification setup. Totally 7 epochs were extracted from 0.5 to 4.0 s after the instruction cue. The epoch length was 2.0 s with 90% overlap between consecutive epochs. EEG signals were band-pass filtered in the selected user-specific frequency band, using a 4^{th} order Butterworth filter.

B. Comparing methods: baseline and conventional method

As baseline, we filtered EEG signals in a broad band 8–30 Hz, encompassing the α and β bands, a typical frequency band used in MI-BCI [1] when no frequency band selection algorithm is applied.

As conventional method, we used the well-known heuristic-based selection algorithm proposed by Blankertz et al. [6]. Note that this method was originally developed to find user-specific frequency bands for the spatial filter-based BCIs, but not for Riemannian ones. This selection algorithm is based on the EEG signals at C3 and C4 channels, after Laplacian spatial filtering. First, the correlation coefficient between the log band-power of each Laplacian channel and its class label is computed. Next, the sum of the absolute values of the correlation coefficient for each frequency is computed over both Laplacian channels. Then, the frequency with the maximum sum of the correlation coefficient is selected as the most discriminant frequency peak. Finally, the edges of the frequency band are extended on each side in 5–35 Hz with 1 Hz frequency resolution. For a more detailed description, the reader can refer to [6, Appendix.A].

C. Experimental pipeline

Each method was applied on the available training set. Once the user-specific frequency band was selected, training and testing data were both filtered with the selected band, and then SCMs were computed for classification. We evaluated the performance of each method by looking at the testing set classification accuracy with an MDRM classifier trained on the training set data.

TABLE I
Classification accuracy for BCI IV dataset 2a [%]

	Baseline	Blankertz et al. [6]	classDis
A01E	**85.81**	80.56	78.87
A02E	52.18	51.88	**54.56**
A03E	**89.78**	85.12	**89.78**
A04E	70.93	**71.13**	65.97
A05E	51.49	79.07	**81.55**
A06E	64.38	61.90	**65.87**
A07E	56.55	58.43	**71.83**
A08E	93.45	**93.95**	93.45
A09E	89.78	**90.28**	89.19
Ave.	72.7 ± 17.3	74.7 ± 14.7	**76.8 ± 13.1**

V. Result and Discussion

Classification accuracies are summarized in Table I. Average accuracy [%] was 72.7 ± 17.3, 74.7 ± 14.7 and 76.8 ± 13.1 for the baseline, the conventional method and our method, respectively. Our method showed the highest average accuracy although no statistically significant differences were observed, probably due to the low subject number.

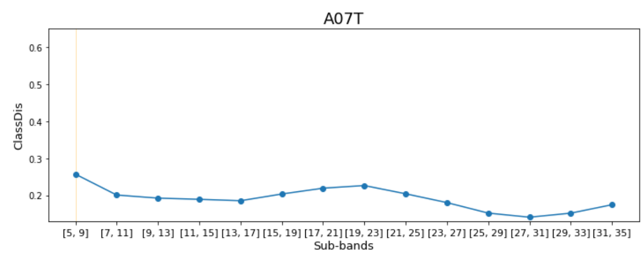

Fig. 1. ClassDis across sub-bands of a subject with large improvement.

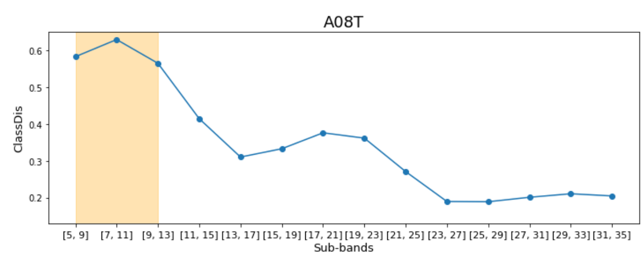

Fig. 2. ClassDis across sub-bands of a subject without improvement.

A. Comparison with baseline

The subjects whose accuracies improved with our method tended to have lower baseline accuracies (ave: $56.2 \pm 5.93\%$) and maximum classDis values (ave: 0.25 ± 0.05). Figure 1 shows the classDis values of one such subject, and the orange shade indicates the selected frequency band for this subject. This may be because those subjects had difficulty in modulating SMRs. Hence, they were our real target. Our method could show substantial improvement for those subjects. Thus, this result suggests the validity of selecting relatively high classDis frequency bands among the overall low classDis bands.

On the other hand, the subjects whose accuracies did not improve with our method tended to have higher baseline accuracy (ave: $86.0 \pm 8.82\%$) and maximum classDis value (ave: 0.61 ± 0.15). Figure 2 shows the classDis values of one of those subjects across all sub-bands. Since baseline accuracies were high, those subjects were expected to be originally good at SMRs modulation. Therefore, they showed overall higher classDis across all sub-bands and selecting frequency bands with even higher classDis might not lead to substantial improvements: the resulting features were already good with a broad band.

B. Comparison with conventional method

There are two main possible reasons that may explain why our method showed a larger average improvement from baseline than the conventional method. The first reason may be the difference in the number of referenced channels for frequency band selection. The conventional method referred only two channels, C3 and C4, whereas our method exploited all channels. Although SMRs modulation are mainly observed in C3 and C4, it may be more effective to take into account the variation of the oscillatory activity globally across all channels to evaluate the class distinctiveness.

The second reason may be due to the consistency between the features used for frequency band selection and for classification. The conventional method used the spectral power of EEG signals as features, while our method used the covariance matrix whose diagonal entries represents the spectral power of each channel. In both cases, the spectral power is taken into account in the frequency band selection. However, there is a major difference between the two methods in the way to evaluate the discriminability of each frequency band. While the conventional method evaluates it by examining the correlation coefficient between the spectral power and the class label at each frequency, our method quantifies the class separability by the ratio of the inter-class distance to the intra-class variance on the manifold. Since MDRM classifies classes based on the distance between classes on the manifold, our method is consistent with the features used for frequency band evaluation and those used for classification. This consistency may be one of the reasons why our method has shown larger substantial improvement, contrary to the conventional method, which is known to improve BCI performance with spectral features [6] - but apparently not necessarily with SCM features.

VI. CONCLUSIONS

In this paper, we have proposed a BCI user-specific frequency band selection method, that works on a Riemannian manifold. The proposed method selects a frequency band on a manifold maximizing class distinctiveness, while adjusting the frequency bandwidth for each individual. In the comparative experiments with a baseline and a conventional method, our method showed the highest average accuracy. In particular, it showed the highest improvements for subjects who originally had low accuracies. This preliminary result suggest the relevance of having a consistent pipeline on a manifold from the frequency band selection to the classification. It also suggests the importance of developing methods based on covariance matrices for Riemannian BCIs, rather than applying conventional methods as they are.

As future works, we will first increase the number of subjects for further evaluation. We tested our method on only a single MI dataset in this paper. However, our class-distinctiveness-based frequency band selection method is generic and should work on any oscillatory activity-based BCI, e.g., on BCI decoding mental calculation or mental workload. Thus, we will also investigate the performance for those datasets. As another future direction, since time segment selection has also shown significant improvements in some literature [19], we are eager to extend our model not only for selecting frequency bands but also relevant time segments for each individual.

REFERENCES

[1] F. Lotte, L. Bougrain, and M. Clerc, "Electroencephalography(EEG)-based Brain-Computer Interfaces," 2015.

[2] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K.-R. Müller, G. Curio, and T. Dickhaus, "Neurophysiological predictor of SMR-based BCI performance," *NeuroImage*, 2010.

[3] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An EEG-based Brain-Computer Interface for cursor control," *EEG and Clinical Neurophys.*, vol. 78-3, pp. 252–259, 1991.

[4] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, "The non-invasive Berlin Brain-Computer Interface: fast acquisition of effective performance in untrained subjects," *NeuroImage*, 2007.

[5] B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio, and K.-R. Müller, "The Berlin Brain-Computer Interface: Accurate performance from first-session in BCI-naive subjects," *IEEE Trans. Biomed. Eng.*, vol. 55-10, pp. 2452–2462, 2008.

[6] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Process. Mag.*, 2007.

[7] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in Brain-Computer Interface," in *Proc. IJCNN*, 2008.

[8] H.-I. Suk and S.-W. Lee, "A novel Bayesian framework for discriminative feature extraction in Brain-Computer Interfaces," *IEEE Trans. Pattern Ana. and Machine Intel.*, vol. 35-2, pp. 286–299, 2012.

[9] C. Benaroch, M. S. Yamamoto, A. Roc, P. Dreyer, C. Jeunet, and F. Lotte, "When should MI-BCI feature optimization include prior knowledge, and which one?," *Brain-Computer Interfaces*, 2022.

[10] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Riemannian geometry applied to BCI classification," in *Int. Conf. Latent Variable Analysis and Signal Separation*, pp. 629–636, 2010.

[11] F. Yger, M. Berar, and F. Lotte, "Riemannian approaches in Brain-Computer Interfaces: a review," *IEEE Trans. Neural Sys. Rehab.*, vol. 25-10, pp. 1753–1762, 2016.

[12] M. Congedo, A. Barachant, and R. Bhatia, "Riemannian geometry for EEG-based Brain-Computer Interfaces; a primer and a review," *Brain-Computer Interfaces*, vol. 4-3, pp. 155–174, 2017.

[13] V. Jayaram and A. Barachant, "MOABB: trustworthy algorithm benchmarking for BCIs," *J. Neural Eng.*, vol. 15-6, p. 066011, 2018.

[14] A. Appriou, A. Cichocki, and F. Lotte, "Modern machine-learning algorithms: for classifying cognitive and affective states from electroencephalography signals," *IEEE Sys. Man. Cybern. Mag.*, vol. 6-3, pp. 29–38, 2020.

[15] X. Pennec, P. Fillard, and N. Ayache, "A Riemannian framework for tensor computing," *Int. J. Comput. Vis.*, vol. 66-1, pp. 41–66, 2006.

[16] F. Lotte and C. Jeunet, "Defining and quantifying users' mental imagery-based BCI skills: a first step," *J. Neural Eng.*, 2018.

[17] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero, "Shrinkage algorithms for MMSE covariance estimation," *IEEE Trans. Sig. Proces.*, vol. 58-10, pp. 5016–5029, 2010.

[18] M. Tangermann, K.-R. Müller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, R. Leeb, C. Mehring, K. J. Miller, G. Mueller-Putz, *et al.*, "Review of the BCI competition IV," *Front. Neurosci.*, 2012.

[19] J. Wang, Z. Feng, X. Ren, N. Lu, J. Luo, and L. Sun, "Feature subset and time segment selection for the classification of EEG data based motor imagery," *Biomed. Sig. Proc. Cont.*, 2020.