PhD proposal:
Similarity-based Classification of EEG

April 2020

1 Riemannian BCI

1.1 Context and references

Brain-computer interfaces (BCI) aim for the real-time decoding of brain signals, users could thus interact with a system without any requirement on their motor abilities. Most BCIs are based on electroencephalography (EEG) as a measure of brain activity. Users with physical disabilities could benefit from EEG-based BCIs with various applications to enhance their quality of life, such as to control spellers, wheelchair or prostheses. This is a trending topic at the interface of neuroscience, signal processing, machine learning and human-machine interface, supported by a large research community and numerous business companies.

To decode brain signals, Riemannian approaches are now the gold standard in many BCI applications [Yger et al., 2016] and have won several data challenges. These approaches rely on the use of spatial covariance matrices as an efficient representation of EEG signals that capture most of the relevant information they contain for classification. The Riemannian framework takes into account the curvature of the space of covariance matrices and yields robust results.

1.2 Limitations

However, Riemannian approaches tend to be inefficient for high dimensional data. To leverage this issue, some Riemannian dimensionality reduction (especially for Symmetric Positive Definite -SPD- matrices) approaches have been developed Harandi et al. [2014], Horev et al. [2015] but they induce an unavoidable loss of information. Using the fact that the submatrices of a SPD matrix have to be SPD, it seems promising to apply the Riemannian geometry on submatrices rather than on the original high dimensional matrix and then to combine those geometries.

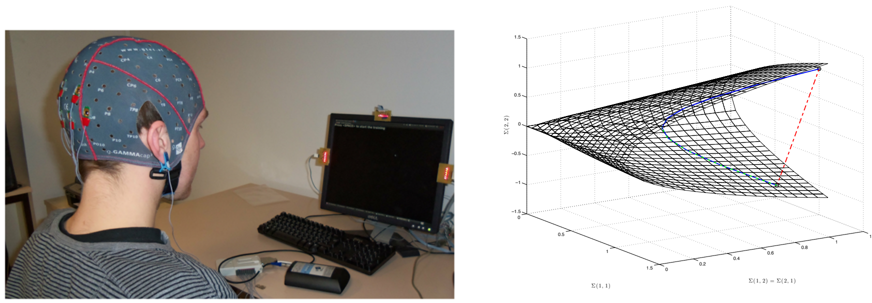

Figure 1: Left: EEG-based brain-computer interface. Right: geodesic and chordal distance for covariance matrices

2 Set-to-Set Metrics

2.1 Scientific objectives of the PhD

When objects are represented as sets of simpler parts, comparing objects can be done using Set-to-Set Metrics. Such an approach has been followed in Gao et al. [2019], where the Set-to-Set metric was learned on top of the Riemannian distance and a dimensionality dimension step. In BCI applications, datasets usually do not contain a lot of samples and due to the number of its parameters, it is unlikely that such an approach could be applied directly without badly overfitting. The ambition of this internship is to propose a simpler version of the approach in Gao et al. [2019] where the Riemannian distance would be fixed, a simpler (and linear instead of quadratic) Set-to-Set Metric would be learned. The dimensionality reduction step would still be learned (and potentially a regularized version in the spirit of Yamane et al. [2016] could be proposed). The developed approach would be then tested on high dimensional BCI dataset with the goal of publishing the obtained results in conferences.

3 Requirements

A good background in applied mathematics and some programming skills in Python are important prerequisites. The work relies on existing open source projects, such as PyRiemann[1] and MOABB[2], to handle datasets and machine learning pipelines. The optimization on manifold will be performed on PyManOpt[3]. The produced code could be released with an open source licence.

[1] `https://github.com/alexandrebarachant/pyRiemann`
[2] `https://github.com/NeuroTechX/moabb`
[3] `https://www.pymanopt.org/`

4 General information

This PhD will be conducted under the supervision of Sylvain Chevallier (LISV) and Florian Yger (LAMSADE). The candidate will be based in the LISV lab, in Velizy. Regular meeting, alternating between LAMSADE and LISV, will take place on a weekly basis.

5 Required qualification and skills

- Second year of Master degree (or last year of engineering school in France) in machine learning, applied mathematics, signal processing or related discipline
- Solid Python programming
- Strong knowledge in machine learning and linear algebra
- Knowledge in EEG and/or BCI a plus but not mandatory
- Knowledge in statistics

6 Main objectives

- Explore and implement (regularized) SPD matrix dimensionality reduction approaches
- Design and implement Set-to-Set metrics, with appropriate metric learning tools, to compute similarity between large SPD matrices, based on (Riemannian) distances between sub-matrices
- Study and validate such approaches offline on real EEG signals for BCI applications
- Extend dimensionality reduction approach for the data visualization, using a simpler and faster approach
- Contribute to existing open source library to disseminate to the obtained results.

7 Supervision

The PhD will be conducted under the direction of Sylvain Chevallier (sylvain.chevallier@uvsq.fr), Associate Professor, LISV (Laboratoire d'Ingéniérie et des Systèmes de Versailles), Versailles, France and of Florian Yger (florian.yger@dauphine.fr), Associate Professor, Univ. Paris Dauphine/LAMSADE, Paris, France.

8 Collaboration and international exchange

Regular interaction are expected with Fabien Lotte (fabien.lotte@inria.fr), Senior Research Scientist, Inria Bordeaux Sud-Ouest (France), to develop the brain-computer interface part.

This research is part of an ongoing collaboration with the Tokyo University of Agriculture and Technology, with the Pr Toshihisa Tanaka (`https://sites.google.com/a/sip.tuat.ac.jp/toshihisa-tanaka/`). An international mobility with Japan is expected, as there is an ongoing demand for international travel grant to support this project.

References

Zhi Gao, Yuwei Wu, Mehrtash Harandi, and Yunde Jia. A robust distance measure for similarity-based classification on the spd manifold. *IEEE Transactions on Neural Networks and Learning Systems*, 2019.

Mehrtash T Harandi, Mathieu Salzmann, and Richard Hartley. From manifold to manifold: Geometry-aware dimensionality reduction for spd matrices. In *European conference on computer vision*, pages 17–32. Springer, 2014.

Inbal Horev, Florian Yger, and Masashi Sugiyama. Geometry-aware principal component analysis for symmetric positive definite matrices. In *ACML*, pages 1–16, 2015.

Ikko Yamane, Florian Yger, Maxime Berar, and Masashi Sugiyama. Multitask principal component analysis. In *Asian Conference on Machine Learning*, pages 302–317, 2016.

Florian Yger, Maxime Berar, and Fabien Lotte. Riemannian approaches in brain-computer interfaces: a review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(10):1753–1762, 2016.

Cette proposition de thèse sera co-encadrée par Florian Yger (LAMSADE, Université Paris-Dauphine `http://www.yger.fr/`) et Sylvain Chevallier (LISV, Université de Versailles Saint-Quentin, Université Paris-Saclay, `https://sylvchev.github.io/`).