Comparison of Feature Extraction Methods for Brain-Computer Interfaces

Patrick Ofner[1], Gernot R. Müller-Putz[1], Christa Neuper[1,2] and Clemens Brunner[1]

[1]*Institute for Knowledge Discovery, BCI Lab, Graz University of Technology, Austria*
[2]*Department of Psychology, University of Graz, Austria*

Introduction

Classification accuracies of feature extraction methods (FEMs) as used in sensory motor rhythm (SMR) based Brain-Computer Interfaces (BCIs) were compared offline. Features were extracted from 9 subjects and classified with linear discriminant analysis (LDA). The following FEMs were compared: adaptive autoregressive parameters (AAR), bilinear AAR (BAAR), multivariate AAR (MVAAR), band power (BP), phase locking value (PLV), time domain parameters (TDP), and Hjorth parameters.

Most FEMs contain meta parameters and it is crucial to tune these meta parameters carefully to tap the full potential of these methods. Therefore, all meta parameters were optimized in a subject-specific way with a genetic algorithm (GA) [1].

Paradigm

In this cue-based paradigm two motor imagery tasks had to be performed: motor imagery (MI) of the left vs right hand. No feedback was provided. Two sessions from different days from each of 9 participants were recorded. One session comprised 6 runs, each with 24 trials. A trial sequence is shown in Figure 1.

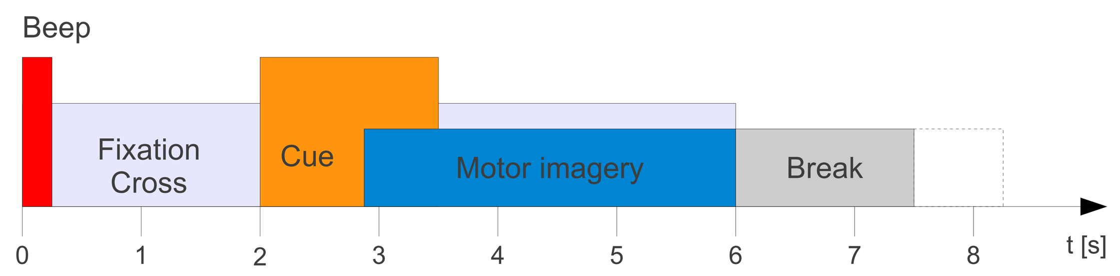

Figure 1: cue-based paradigm

Test Setup

FEMs were optimized in the *optimization step* with a genetic algorithm, and afterwards tested in the *evaluation step*. Data from session 2 was used solely for testing the LDA classifier in the evaluation step. See Figure 2.

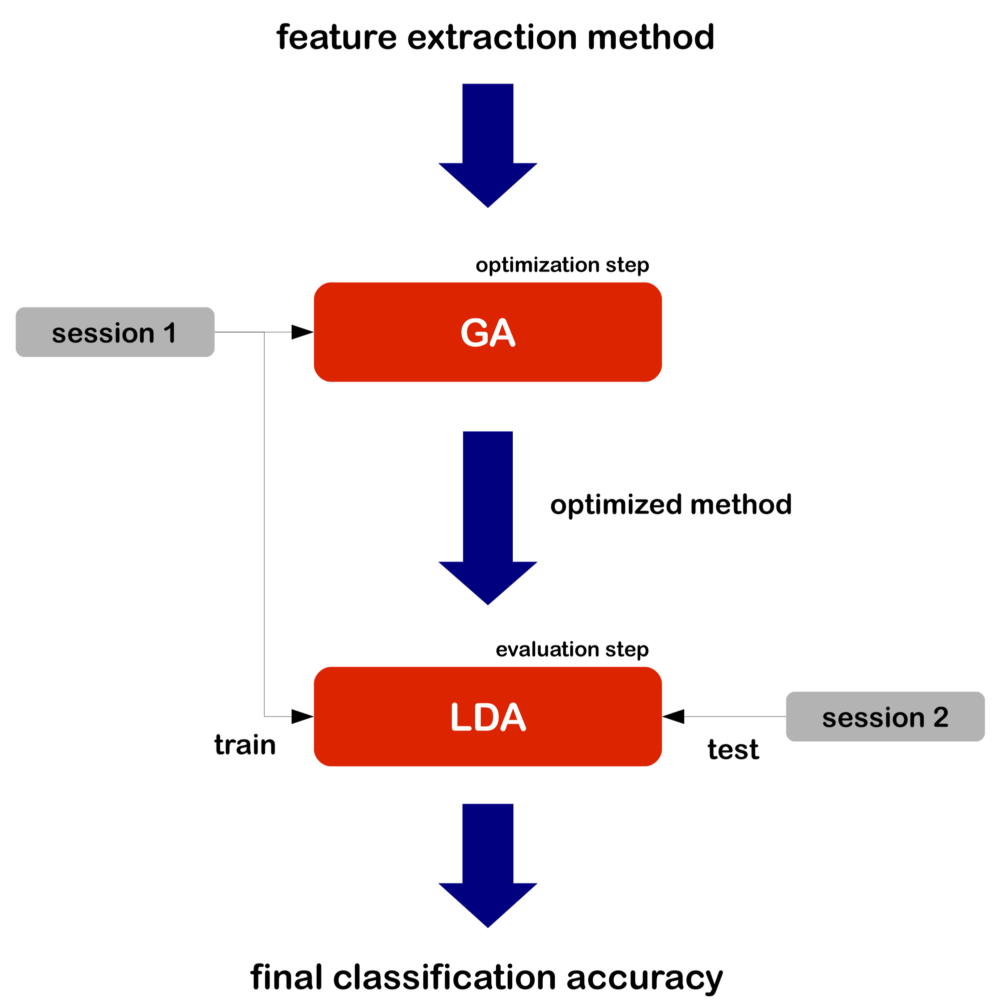

Figure 2: procedure for determining the classification accuracy of a feature extraction method

All FEMs used data from C3 and C4 positions with a specific spatial filter (monopolar, bipolar, Laplace, common average reference (CAR)). Thereby, three types of bipolar spatial filters were used: FC3/4-C3/4, C3/4-CP3/4, FC3/4-CP3/4. The type of bipolar spatial filter with the best fitness score (best classification accuracy) in the optimization step for a subject was used. In addition, the PLV FEM used four channels (two channels per hemisphere) in various arrangements, because inter-hemispheric coupling was expected to contain discriminative information [2]. The channel combination leading to the best fitness score in the optimization step was used for further analysis.

In the evaluation step features were extracted using the optimized meta parameters. An LDA classifier was trained with features from session 1 of a subject and tested against features from session 2 of the same subject. The 0.9 quantile of the classification accuracy reached by the LDA classifier (between cue and end of trial) was reported as the classification accuracy for a FEM.

Results

Figure 3 shows on the left side a box-and-whisker plot including mean values (dotted lines). Only spatial filters yielding the highest mean classification accuracy are shown. TDP with a bipolar spatial filter reaches the highest *mean* classification accuracy of 78 % (standard deviation 11 %, median 82 %). MVAAR with a bipolar spatial filter reaches the highest *median* classification accuracy of 83 % (standard deviation 13 %, mean 74 %). On the right side, Figure 3 shows mean values and standard deviations of all FEMs and spatial filters.

Tukey's Test has been used to test for significant differences ($\alpha = 0.05$) of the FEMs shown in Figure 3 and spatial filters when using TDP. PLV (CAR) differs significantly from TDP (bipolar) and BP (bipolar). A monopolar spatial filter is significantly worse than bipolar and Laplacian filters when using the method with the highest mean classification accuracy (TDP).

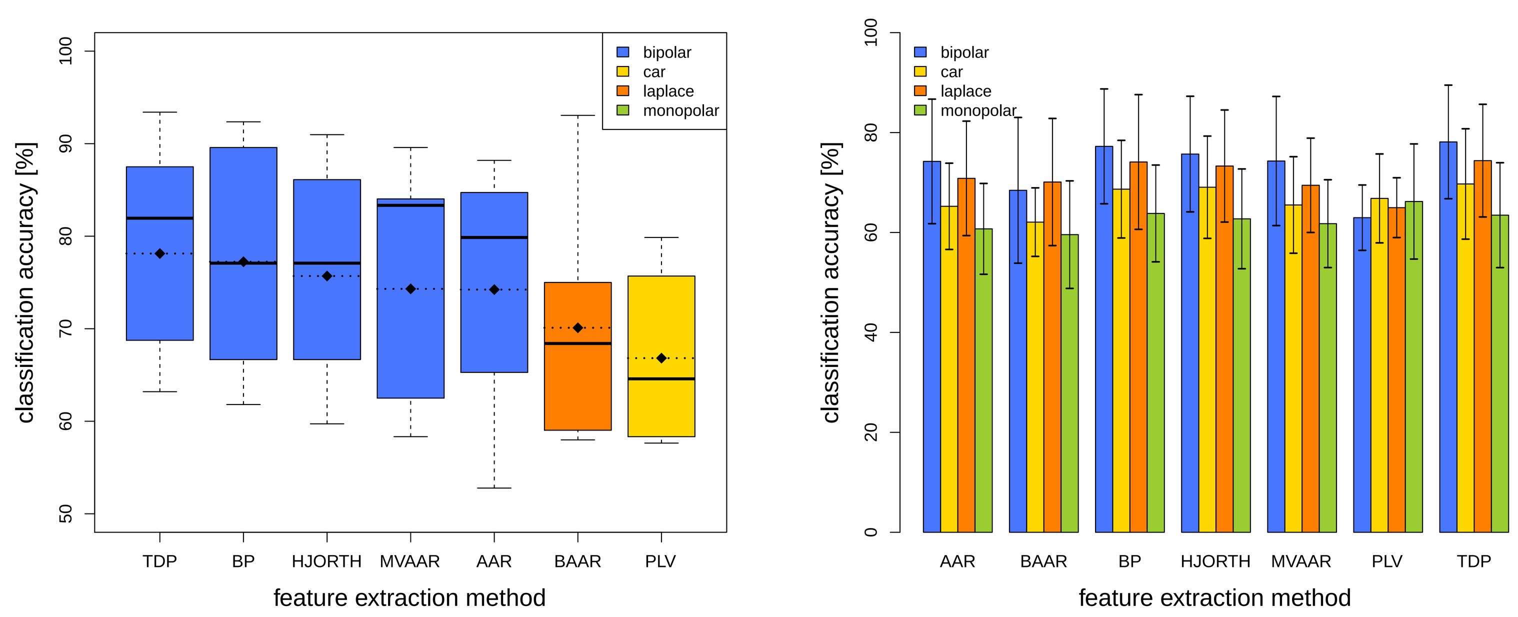

Figure 3: left: box-and-whisker plot including mean values, right: mean values and standard deviations of all FEMs and spatial filters

Conclusion

No significant differences were found between TDP, BP, Hjorth, MVAAR, AAR, BAAR. However, TDP with a bipolar spatial filter yielded the highest mean classification accuracy, a high median classification accuracy, is computationally efficient, has less parameters to set. TDP is therefore favorable of all compared feature extraction methods when using an LDA classifier, a small number of electrodes and a comparable paradigm.

References

1. J. Holland, Adaptation in natural and artificial systems, *University of Michigan Press (reprinted in 1992 by The MIT Press)*, 1975
2. C. Brunner, R. Scherer, B. Graimann, G. Supp, and G. Pfurtscheller, Online control of a brain-computer interface using phase synchronization, *IEEE Transactions on Biomedical Engineering*, 53(12):2501–2506, 2006

Acknowledgments

This work is supported by the European ICT Programme Project FP7-224631. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.