# Using random forests for classifying motor imagery EEG

D. Steyrl[1], R. Scherer[1], G. R. Müller-Putz[1]

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

## Introduction

One crucial issue for accurate sensorimotor rhythm (SMR) Brain-Computer Interfaces (BCIs) control is the selection of the most discriminative oscillatory EEG components. The Random Forests (RF) ensemble classifier is particularly interesting in this context because he:

- can handle high dimensional input variables [1].
- achieves high classification accuracies and is robust against outliers [1].
- has a built in feature rating and selection mechanism (Gini Index)[1].

We performed off-line analysis of right hand vs. feet MI data and compared the feature rating results with event-related desynchronization /synchronization (ERD/S) maps [2].

## Methods

**Random Forests classifier and Gini Index:** A RF classifier is an ensemble of many decision trees (Fig. 1).

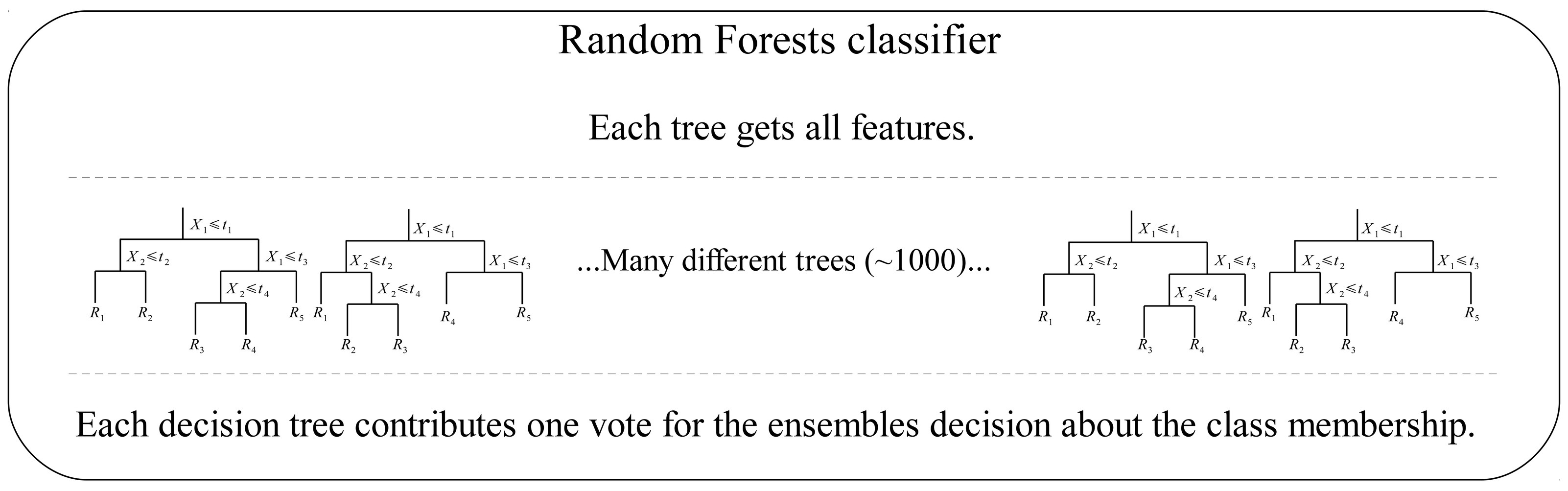


**Figure 1:** Structure of an RF classifier.

An RF classifier's accuracy depends on two things [1]:

- Low correlation between the trees.
- High accuracy of the individual trees.

The correlation can be decreased with randomness:

- Through an individual bootstrap sample for each tree.
- Through an random feature subset for each split in each tree.

The tree's accuracy is high if the separation of inputs is pure. The purity is measured with the Gini Index (GI) [3]. If the GI is decreased by a feature means that this feature contains information about the class membership and is therefore important [1].

**The Paradigm is shown in Fig. 2:**

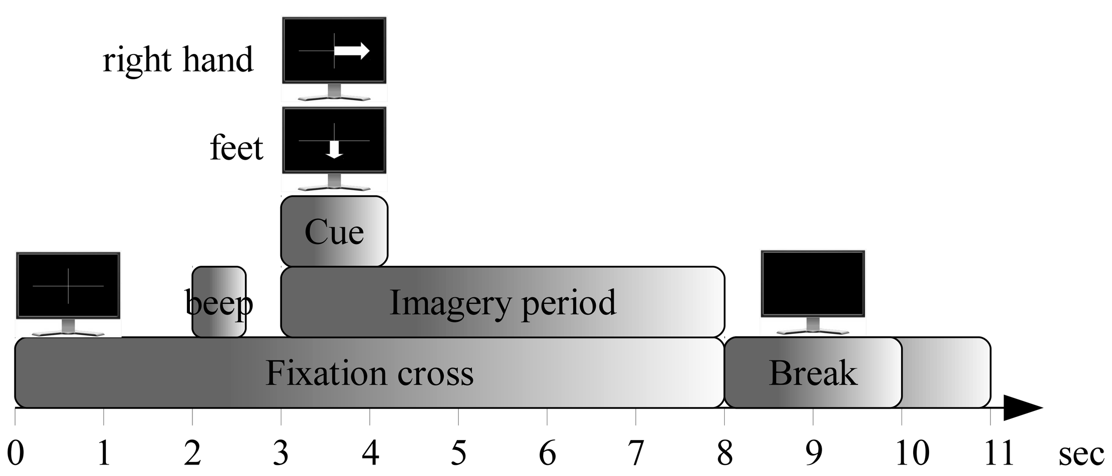


**Figure 2:** EEG recorded by a standard cue-based paradigm with 4 s imagination periods of right hand and feet motor executions was analyzed [4]

**Signal processing:**

- 10 datasets, recorded for [4].
- Laplacian derivations of C3, Cz and C4 (10-20 system).
- FFT on overlapping (0.5 s) windows with a length of 1 s.
- Absolute values of the frequency range from 5 to 40 Hz at a frequency resolution of 1 Hz.
- Artifact free trials (visual inspection), to calculate 10x10 fold cross-validation (CV) accuracies for each time window of each participant independently.
- Correlation coefficients between GI based feature ratings and ERD/S maps [2].

## Results

The CV results and the correlation coefficients of each participant are presented in Tab 1. Fig. 3 shows one example and the average GI feature rating over all participants.

**Table 1:** Peak CV accuracies of RF classifiers and corr. coeff. between RF classifiers feature rating and significant (99% confidence interval) differences of ERD/S maps of the classes right hand vs. feet (n.s. not significant).

| Participant | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | Mean |
|---|---|---|---|---|---|---|---|---|---|---|---|
| RF | 94% | 70% | 75% | 86% | 91% | 64% | 70% | 85% | 74% | 80% | 79% |
| Corr. coeff. (p <0.001) | 0.45 | n.s. | n.s. | 0.34 | 0.16 | 0.44 | n.s. | 0.17 | 0.12 | 0.71 | 0.34 |

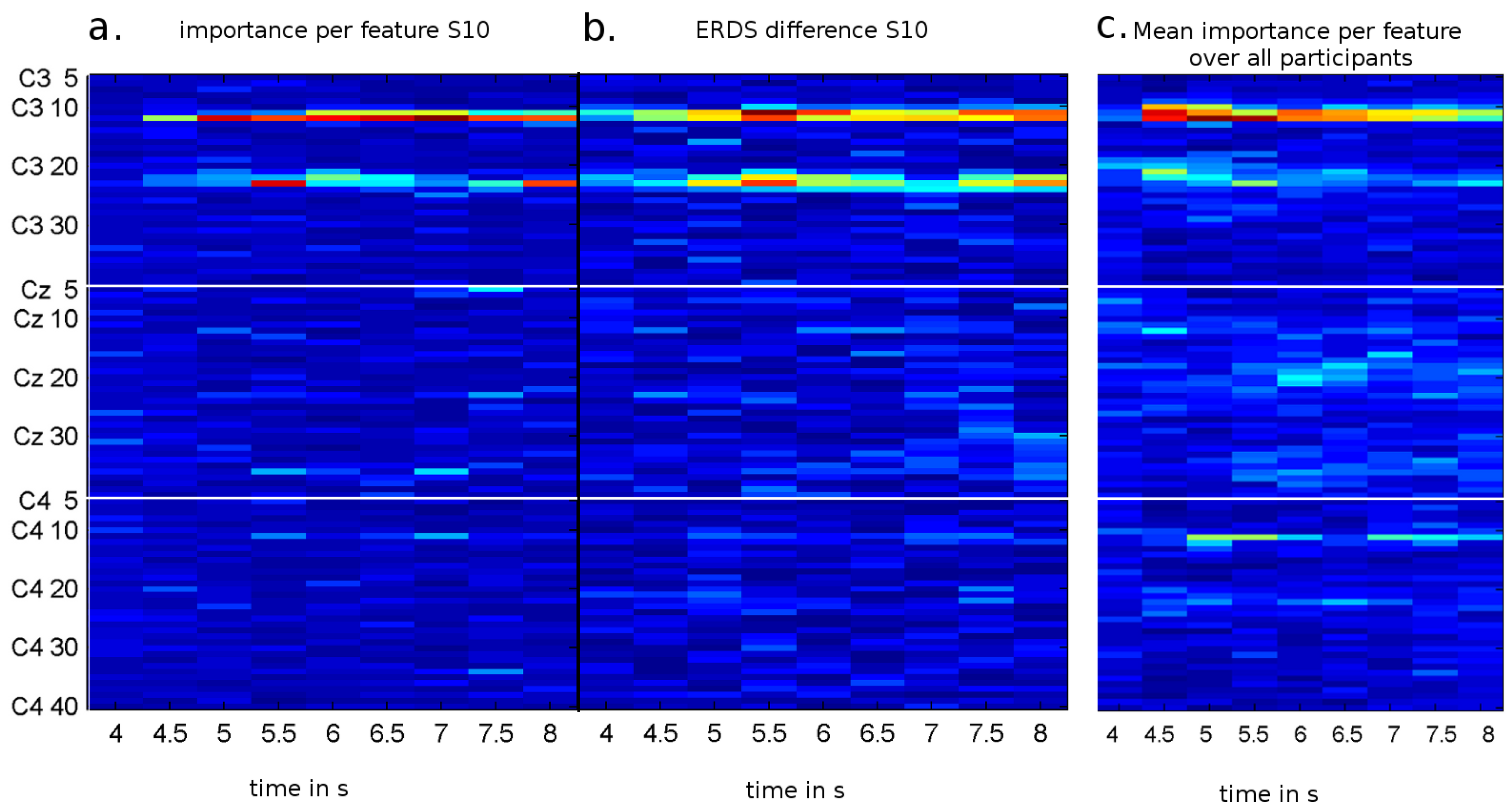


**Figure 3:** 3a) GI rating map for right hand vs. feet MI of participant S10. Each time segment was individually analysed. 3b) difference of ERD/S time/frequency maps between right hand and feet MI for participant S10. 3c) average GI ranking average over all participants. Note: Color coding of the maps is not compareable (normalized to the respective map).

## Discussion

- We computed an average peak accuracy of 79%. For comparison DSLVQ achieved 81% [4].
- Fast training (about 1 s per classifier).
- Fast classification (about 0.01 s per sample), possible on-line application.
- The top rated features for MI were in average in the mu and in the beta band of C3 (Fig. 3 c), which is in line with literature [2].
- Correlation coefficients (Tab. 1) between GI rankings (Fig. 3 a) and differences of ERD/S maps (Fig. 2 b) were low.

Summing up, Random Forests classifier find neurophysiological reasonable features for classifying right hand vs. feet MI data.

## References

1. Breiman L, Random Forests, *Kluwer Academic Publishers, Machine Learning*, 45: 5-32, 2001.
2. Pfurtscheller G. and Lopes da Silva F. H., Event-related EEG/MEG synchronization and desynchronization: basic principles, *Clinical Neurophysiology*, 110: 1842-1857, 1999.
3. Breiman L, Friedman J H, Olshen R A, Stone C J, CART: Classification and Regression Trees *Wadsworth: Belmont, CA*, 1983.
4. Müller-Putz G R, Scherer R, Pfurtscheller G, Neuper C, Temporal coding of brain patterns for direct limb control in humans. *Frontiers in Neuroscience*: 4-34, 2010.