BCI Competition IV, Dataset 1: Motor Imagery, Uncued Classifier Application

Cédric Gouy-Pailler[1]*, Jérémie Mattout[2], Marco Congedo[1] and Christian Jutten[1]

1- GIPSA-lab - Images and Signal Department - CNRS UMR 5216
Joseph Fourier University - Grenoble Institute of Technology - Stendhal University
Domaine Universitaire, BP 46, 38402 Saint Martin d'Hères Cedex - FRANCE
cedric.gouypailler@gipsa-lab.inpg.fr

2- INSERM U821 - Dynamique Cérébrale et Cognition
Institut Fédératif des Neurosciences - Université Lyon 1
Centre Hospitalier Le Vinatier, 95 Boulevard Pinel, 69500 Bron - FRANCE

1 Calibration Data: Offline Training of parameters

1.1 Spatial Filtering

Common Spatial Pattern by Approximate Joint Diagonalization [1] is performed on the calibration data (only class 1 versus class -1, covariance matrices are computed using each EEG channel from $t = 0$ to $t = 4\,s$ of each trial). Among the resulting spatial filters we select the best 10 filters.

1.2 Features Extraction

The features are the log of the energy of the 10 components in 9 different frequency bands (7-9, 9-11, 11-13, 13-15, 15-17, 17-19, 19-21, 21-23, 23-25 Hz). The frequency-specific signals are obtained using 5-order IIR butterworth filters. This yields a total of 90 features for each trial.

1.3 Classification

The classification is performed using a regularized version of logistic regression (elasticnet regularization, class 1 versus class -1) [2]. The code is available in the glmnet package (CRAN R open-source software). A weight is applied on every trials to give more importance to the most recent trials.

1.4 Other Parameters

The calibration data is also used to set the following parameters by cross-validation:

- Best window length for classification among 1 s, 1.5or 2 s: for each subject we finally use 1.5 s;

*This work was supported by the French Ministry of Defense (Délégation Générale pour l'Armement – DGA), the Open-ViBE project, and the European program of Cooperation in the Field of Scientific and Technical Research (COST B27).

- Best window position: this parameter is subject-dependent;
- Classifier regularization parameter: among a set of determined values, we find that $\lambda = 0.05$ yields the best crossvalidation results.

2 Evaluation Data

In the evaluation step we do not have any cue, we thus consider each successive overlapping time windows for classification (window length is 1.5 s and the overlapping is 80%).

2.1 Online Adaptive Classification

Features of each clearly classified segment are added to the constantly updating training set (fixed size 100 trials for class 1 and 100 trials for class -1). The same weights, aiming at giving more importance to recent trials and applied to the updated training set, is used to retrained the classifier every 30 s.

The classification of each segments yields a probability of being in class 1 or -1. We decide to use a relax class when the classifier cannot clearly decide between class 1 and -1. Thus, by defining $\mu = \alpha|p(-1|\text{imagery}) - p(1|\text{imagery})|$, we have

$$\begin{aligned} p(\text{rest}) &= 1 - \mu & (1) \\ p(1) &= \mu.p(1|\text{imagery}) & (2) \\ p(-1) &= \mu.p(-1|\text{imagery}) \ . & (3) \end{aligned}$$

α=0.9 works fine for every subject.

2.2 Mental States Transition Filtering by Variational HMM

The three previous probabilities are lastly filtered by a variational Bayesian HMM [3] to constrain the Mental states transition dynamics. The hidden state is a three-state variable (rest, -1 or 1), the observations are the probabilities. A priori probability transitions are chosen according to the expected imagery duration (from 1.5 to 5 seconds).

We provide the predicted class (class with maximum probability for each window). The results is then resampled at 100 Hz to get one result at each time point. The resampling is made by replicating 30 times each results, and padding with 0 at the beginning and the end.

3 Note on Artificially Generated Datasets

We think that datasets e, f (and maybe g) are artificial data: it seems that a natural shift of probabilities between calibration and evaluation data is observed in real data but might not be modeled in artificial datasets (at least it is not observed in the mentioned datasets).

References

[1] Cédric Gouy-Pailler, Marco Congedo, Christian Jutten, Clemens Brunner, and Gert Pfurtscheller. Model-based source separation for multi-class motor imagery. In *Proceedings of the 16th European Signal Processing Conference (EUSIPCO-2008), EURASIP*, Lausanne, Switzerland, August 2008.

[2] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. Regularized paths for generalized linear models via coordinate descent. The R package glmnet is available from CRAN, 2008.

[3] Václav Šmídl and Anthony Quinn. *The Variational Bayes Method in Signal Processing (Signals and Communication Technology)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.