Optimizing Spatial Filters for BCI

Optimizing Spatial Filters for BCI

J. Farquhar, N. J. Hill, B. Schölkopf
e-mail: {jdrf,jez,bs}@tuebingen.mpg.de
Max Planck Institute for Biological Cybernetics, Tübingen, Germany

INTRODUCTION

We present easy-to-use alternatives to the often-used two-stage Common Spatial Pattern [1]+classifier approach for spatial filtering and classification of Event-Related Desychnronization signals in BCI. We report two algorithms that aim to optimize the spatial filters according to a criterion more directly related to the ability of the algorithms to generalize to unseen data. Both are based upon the idea of treating the spatial filter coefficients as hyperparameters of a kernel or covariance function. We then optimize these hyper-parameters directly along side the normal classifier parameters with respect to our chosen learning objective function. The two objectives considered are margin-maximization as used in Support-Vector Machines [2], and the evidence maximization framework used in Gaussian Processes [3].

RESULTS

Preliminary results below show average generalization error over 8 test folds, on 5 offline motor imagery data sets measured in Tübingen. Both the our approaches show consistent improvements relative to the commonly used CSP+linear classifier combination. Strikingly, the improvement is most significant in the higher noise cases, when either few trails are used for training, or with the most poorly performing subjects.
This a reversal of the usual "rich get richer" effect in the development of CSP extensions (such as CSSP [4] or CSSSP [5]) which tend to perform best when the signal is strong enough to accurately find their additional parameters. This makes our approach particularly suitable for clinical application where high levels of noise are to be expected.

	100/300 (Train/Test)					200/200 (Train/Test)				
Subj	hm	je	jv	ms	nl	hm	je	jv	ms	nl
CSP	34	24	10	02	45	29	21	09	02	34
MM	**27**	20	**05**	**01**	**37**	**24**	18	**05**	**01**	**30**
GP	28	**19**	**05**	02	**37**	26	**16**	05	02	32

Table 1 Error rates (%) for the different algorithms.

RERFERENCES

[1] Koles Z. J, Lazar M. S, and Zhou S. Z. Spatial patterns underlying population differences in the background EEG. *Brain Topography* **2**(4), 275–284 (1990).

[2] Schölkopf, B. and A.J. Smola *Learning with Kernels* MIT Press, (2002).

[3] Rasmussen C. E., and C.K.I. Williams *Gaussian Processes for Machine Learning* MIT Press, Cambridge, MA (2006).

[4] Lemm S, Blankertz B, Curio G, and Müller K.-R. Spatio-spectral filters for robust classification of single trial EEG. *IEEE Trans. Biomedical Eng.* **52**(9), (2004).

[5] Dornhege G, Blankertz B, Krauledat M, Losch F, Curio G, and Müller K.-R. Optimizing spatio-temporal filters for improving brain-computer interfacing. In: Weiss Y, Schölkopf B, and Platt J (Eds.), *Advances in Neural Information Processing Systems 18*. MIT Press, Cambridge, MA (2006).