A Bayesian Approach for Adaptive BCI Classification

M. Kawanabe[1], M. Krauledat[1,2], B. Blankertz[1]

[1]Fraunhofer FIRST (IDA), Berlin, Germany
[2]Technical University Berlin, Berlin, Germany
E-mail: nabe@first.fhg.de

SUMMARY: In this article, we present an adaptive classifier for BCI based on a mixture of Gaussian (moG) model of the features and a dynamical Bayesian model of the class means. We apply this approach to feedback data from the Berlin Brain-Computer Interface (BBCI). The proposed model can improve the classification performance by compensating for substantial changes of EEG signals between training and feedback sessions as well as for gradual nonstationarity in the feedback sessions.

INTRODUCTION

EEG-based BCI systems are often subject to nonstationarities that are caused by changes in the subject's mental state during an experiment (e.g. due to fatigue, change of task involvement and demands for visual processing etc.). Recently Shenoy et al. [2] showed that a simple bias recalculation for the classifier obtained from the training data can eliminate sources most detrimental effects of nonstationarities during feedback operation. In this paper, we propose a Bayesian version of such adaptive classifiers, where the class means are treated as random variables and their posterior distributions are approximated by a sequential manner as Kalman filters. The proposed method was applied to BBCI data collected from three subjects.

MATERIALS AND METHODS

We investigate data from a study of three subjects using the BBCI system similar to [1] but with very long feedback blocks without break. The experiments consisted of a calibration measurement and a feedback period. In the calibration measurement, visual stimuli (L), (R) (for imagined left and right hand movement) and (F) (for imagined foot movement) were presented to the subjects. Based on the recorded signals, subject-specific features for the further analysis were calculated. The most discriminative frequency band for two of the three classes was selected manually by experts, and common spatial patterns (CSP) were calculated. For the data sets we analyzed, 6 (*al*), 2 (*aw*) and 4 (*VPt*) CSP channels were used, respectively. The bandpower of the CSP-projected channels was estimated using windows of 3 seconds length, and finally a linear classifier was trained by linear discriminant analysis (LDA).

In the feedback phase, bandpower estimations of CSP channels were calculated in a similar manner as in the calibration session for sliding windows of 1 second length. The real-valued output by the LDA classifier was used to move a cursor horizontally on the screen. The subjects were then using this cursor for the operation of a text input (speller) software.

We employ the following mixture of Gaussian (moG) model for each class distribution

$$\begin{aligned} p(\boldsymbol{x}|y=1) &:= (1-p_{\mathrm{p}})\phi(\boldsymbol{x}|\boldsymbol{\mu}_{\mathrm{p}},\Sigma_{\mathrm{p}}) + p_{\mathrm{p}}\phi(\boldsymbol{x}|\boldsymbol{m},V), \\ p(\boldsymbol{x}|y=-1) &:= (1-p_{\mathrm{n}})\phi(\boldsymbol{x}|\boldsymbol{\mu}_{\mathrm{n}},\Sigma_{\mathrm{n}}) + p_{\mathrm{n}}\phi(\boldsymbol{x}|\boldsymbol{m},V), \end{aligned}$$

where $\phi(\boldsymbol{x}|\boldsymbol{\mu},\Sigma)$ is the Gaussian density function with mean $\boldsymbol{\mu}$ and covariance Σ. The first terms represent typical samples, while the common second term corresponds to outliers with large covariance V. Although we concentrated on the binary classification problem, the moG model also enables us to recognize outlying observations from typical samples. In the training session, we estimate the model parameters, i.e., the mean ($\boldsymbol{\mu}_{\mathrm{p}}$, $\boldsymbol{\mu}_{\mathrm{n}}$, $\boldsymbol{m}$) and the covariance (Σ_{p}, Σ_{n}, V) of each Gaussian prototype, their outlier ratios (p_{p}, p_{n}), and the class probability ($\pi := p(y=1)$) by EM algorithm with an extra restriction to keep the covariance V of the outlier large.

To cope with the difference of EEG signals between training and feedback and the gradual nonstationarity in the feedback session we assume that the centers of both classes are random variables and subject to the dynamical model ($t \geq 1$)

$$\begin{aligned} \boldsymbol{\mu}_{\mathrm{p}}(t) &= \boldsymbol{\mu}_{\mathrm{p}}(t-1) + \boldsymbol{\varepsilon}_{\mathrm{p}}(t), \\ \boldsymbol{\mu}_{\mathrm{n}}(t) &= \boldsymbol{\mu}_{\mathrm{n}}(t-1) + \boldsymbol{\varepsilon}_{\mathrm{n}}(t), \end{aligned}$$

where $\boldsymbol{\varepsilon}_{\mathrm{p}}(t) \sim N(\mathbf{0},\Delta_{\mathrm{p}})$, $\boldsymbol{\varepsilon}_{\mathrm{n}}(t) \sim N(\mathbf{0},\Delta_{\mathrm{n}})$. The initial means are also assumed to be Gaussians centered at the estimators from the training session, i.e. $\boldsymbol{\mu}_{\mathrm{p}}(0) \sim N(\widehat{\boldsymbol{\mu}}_{\mathrm{p}},\Gamma_{\mathrm{p}})$ and $\boldsymbol{\mu}_{\mathrm{n}}(0) \sim N(\widehat{\boldsymbol{\mu}}_{\mathrm{n}},\Gamma_{\mathrm{n}})$, respectively. The covariances Δ_{p}, Δ_{n}, Γ_{p} and Γ_{n} control the speed of adaptation and should be chosen according to the magnitude of the initial covariances. The center $\boldsymbol{m}(t)$ of the outlier class is fixed at the average of the positive and negative classes. The required parameters are determined on the training data.

When the samples and the labels $\mathcal{D}_t = \{\boldsymbol{x}_\tau, y_\tau\}_{\tau=1}^{t}$ up to t-th trial are observed, we infer the posterior distribution $p(\boldsymbol{\mu}_{\mathrm{p}}(t), \boldsymbol{\mu}_{\mathrm{n}}(t)|\mathcal{D}_t)$ by a sequential scheme as Kalman filter. However in contrast to the case in Kalman filters, the posterior is not Gaussian in our moG model. Hence, we approximate it by a single Gaussian distribution with the same mean and covariance. We construct a classifier based on the posterior distribution in order to predict the label of the $(t+1)$th trial from the inputs. In this study we adopted the classifier based on the posterior probability of the typical positive minus that of the typical negative class, i.e. $f_t(\boldsymbol{x}) := P(y = 1, z = 0|\boldsymbol{x}, \mathcal{D}_t) - P(y = -1, z = 0|\boldsymbol{x}, \mathcal{D}_t)$, where the latent variable z equals 1 if the sample is an outlier and 0 otherwise.

RESULTS

In our Bayesian framework, the posterior distributions of the class means $\boldsymbol{\mu}_{\mathrm{p}}(t)$ and $\boldsymbol{\mu}_{\mathrm{n}}(t)$ are approximated by Gaussians. In order to visualize the nonstationarity of the data infered by the moG model, we plotted the time course of the posterior means in Figure 1, where the horizontal axis is the direction of the original classifier. Time is indicated by gray scale (black to white). At the beginning, because less information about the class means $\boldsymbol{\mu}_{\mathrm{p}}(t)$ and $\boldsymbol{\mu}_{\mathrm{n}}(t)$ is available, the posterior means can move by a large amount, while the changes get smaller as more trials are performed. For the subject *aw*, although the mean of the positive class again comes closer to the estimator from the training data after feedback learning, that of the negative class seems to stay away from the original estimator. This is the reason why classifier modification in the feedback session can improve the performance.

In Table 1, we compare the classification error in the feedback sessions of our approach (ADB) with the error of the original classifier (ORIG). ADB was much better for subject *aw*, equal for *VPt* and worse for *al*.

Table 1: Comparison of classification errors (window-wise, the feedback error in the BCI task was lower)

	ORIG	ADB
al	10.5	13.4
aw	22.5	9.0
VPt	18.5	18.3

DISCUSSION

The number of data sets limits the interpretation of our results so we can only speculate. In line with the results of [2], one possible reason for the difference in relative performance could be that the features which are extracted by the BBCI are often not strongly affected by nonstationarities (*al*, *VPt*). Accordingly adaptive methods can improve only for some datasets (*aw*). Another observation is that the in/decrease of performance correlates with the number of features. It could be that some parameters in the more complex method ADB are not accurately estimated. This issue and possible remedies are subject to further research.

REFERENCES

[1] Blankertz B, Dornhege G, Krauledat M, Müller K-R, Kunzmann V, Losch F, Curio G. The Berlin Brain-Computer Interface: EEG-based communication without subject training. *IEEE TNSRE*, 14(2), 2006 (in press).
[2] Shenoy P, Krauledat M, Blankertz B, Rao RPN, Müller KR. Towards adaptive classification for BCI. *J. Neural Eng.*, 3:R13–R23, 2006.

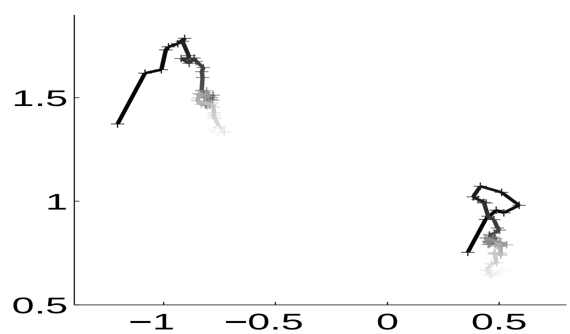

Figure 1: Time course of the class means. The posterior means of the random variables $\boldsymbol{\mu}_{\mathrm{p}}(t)$ and $\boldsymbol{\mu}_{\mathrm{n}}(t)$ for the subject *aw* were plotted so that the horizontal axis coincides with the direction of the original classifier.