Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

Adaptive Classification for Brain Computer Interfaces

Julie Blumberg, Jörn Rickert, Stephan Waldert, Andreas Schulze-Bonhage, Ad Aertsen, Carsten Mehring

Abstract— **In this paper we evaluate the performance of a new adaptive classifier for the use within a Brain Computer-Interface (BCI). The classifier can either be adaptive in a completely unsupervised manner or using unsupervised adaptation in conjunction with a neuronal evaluation signal to improve adaptation. The first variant, termed Adaptive Linear Discriminant Analysis (ALDA), updates mean values as well as covariances of the class distributions continuously in time. In simulated as well as experimental data ALDA substantially outperforms the non-adaptive LDA. The second variant, termed Adaptive Linear Discriminant Analysis with Error Correction (ALDEC), extends the unsupervised algorithm with an additional independent neuronal evaluation signal. Such a signal could be an error related potential which indicates when the decoder did not classify correctly. When the mean values of the class distributions circle around each other or even cross their way, ALDEC can yield a substantially better adaptation than ALDA depending on the reliability of the error signal. Given the non-stationarity of EEG signals during BCI control our approach might strongly improve the precision and the time needed to gain accurate control in future BCI applications.**

I. INTRODUCTION

Brain Computer Interfaces (BCIs) infer the movement intentions of subjects by analyzing electrophysiological signals of the brain [1]. A prominent approach to BCIs uses a supervised training phase where brain signals as well as movement imagery are recorded and a classifier is trained to separate different movements, e.g. left or right hand movement [2]. Later, in the feedback mode with direct online control, the classifier uses the current brain activity to estimate the movement intention. The classification outcome is presented to the subject for instance as a cursor movement on a screen. However, the recorded signals can undergo considerable changes between training and feedback mode as well as during feedback itself [3], [4], [5]. Such non-stationarities in the signals can be due to

- task differences between training and feedback, e.g. real vs. imagined actions,
- variability of the recording caused by drying gel or micro movements of the electrodes,
- plasticity of the brain, due to experience with the task,
- modulation of cognitive states like attention, motivation and vigilance [3].

This study was supported by the German Federal Ministry of Education and Research (BMBF grant 01GQ0420 and 01GQ0761 TP 1 and the GoBio grant 0313891).

J. Blumberg, J. Rickert, S. Waldert, A. Aertsen, and C. Mehring are with the Department of Neurobiology & Bernstein Center for Computational Neuroscience, University of Freiburg, 79104 Freiburg, Germany.

A. Schulze-Bonhage and J. Blumberg are with the Epilepsy Center of the University Hospital, Albert-Ludwigs-University Freiburg, Germany.

Contact: blumberg@biologie.uni-freiburg.de

The last two causes are due to properties of the human brain and can thus hardly be avoided. The readout of movement intentions of a BCI user therefore benefits from adaptive decoding algorithms which are able to track the brain signal changes. As revealed in [3], adaptive methods can significantly increase the decoding performance in the case of non-stationarities, but most adaptive algorithms suggested so far (e.g. [3], [4], [6], [7]) are based on supervised learning techniques, which require the knowledge of the true movement intention during the adaptation. For a BCI application it would be highly desirable to use an algorithm that can interpret the movement intention of the user in an unsupervised manner, because this allows the persons to interact with their environment without the need to constantly reaffirm their intended motion in an external modality. In addition, neuronal evaluation signals, like error related negativity, might be used to improve adaptation. Here, we propose such an adaptive BCI decoding approach by an extension of the widely used Linear Discriminant Analysis classification algorithm.

II. METHODS

A. Linear Discriminant Analysis (LDA)

A commonly used classification algorithm in BCI research is the Linear Discriminant Analysis (LDA). A linear classifier in general tries to establish a hyperplane separating the signal space into individual subspaces for all classes. In the binary case, the decision rule for a given vector x to belong to class $c1$ and not $c2$ reads:

$$p(c_1|x) > p(c_2|x), \tag{1}$$

These probabilities can be computed using Bayes' formula:

$$p(c_k|x) = \frac{p(c_k)p(x|c_k)}{p(x)} \tag{2}$$

where $p(c_k)$ is the the prior probability for a class k and $p(x|c_k)$ is the class distribution. Assuming that all classes are a priori equally probable, the priors can be neglected here. Therefore, the decision rule reduces to

$$p(x|c_1) > p(x|c_2). \tag{3}$$

In the case of LDA, a multivariate Gaussian distribution is assumed for each of the classes c_k,

$$p(x|c_k) = \frac{1}{\sqrt{(2\pi)^f \det(C)}} \exp(-\frac{1}{2}(x-\mu_k)^T C^{-1} (x-\mu_k)) \tag{4}$$

1-4244-0788-5/07/$20.00 ©2007 IEEE

where x is the vector to be classified, f is the dimension of this vector, C is the common covariance matrix for all classes and μ_k is the mean value of class k. In this case the decision surface is given by a hyperplane separating the signal space in two subspaces.

The classification performance of a classifier can be measured by the *decoding power*, i.e. the percentage of correctly classified samples.

B. Adaptive Linear Discriminant Analysis (ALDA)

In brief, the adaptive LDA (ALDA) works as follows: Using data from a supervised training period like for the LDA, the initial data distributions in feature space are estimated for the different classes. In every following feedback step the discriminative condition is then updated via Expectation-Maximization [8], where

- In the Expectation step, the current probability is estimated using Gaussian distributions for K classes:

$$p(c_k|x) = \frac{p(c_k)p(x|c_k)}{p(x)} = \frac{p(c_k)p(x|c_k)}{\sum_{k=1}^{K} p(x|c_k)p(c_k)} \quad (5)$$

- In the Maximization step, the resulting probability of the expectation is used to update the means for all classes, calculated of the last N vectors x_i weighted with their class probabilities (for training data, where a class label k is known, the probability for the correct class is 1, for all other classes 0),

$$\mu_k = \frac{1}{p(c_k)N} \sum_{i=1}^{N} p(c_k|x_i)x_i \quad (6)$$

  and to update the common covariance

$$C = \frac{1}{N-1} \sum_{k=1}^{K} \sum_{i=1}^{N} p(c_k|x_i)(x_i - \mu_k)(x_i - \mu_k)^T. \quad (7)$$

For every new sample and its estimated probabilities, the oldest of a N-sample window of the former training samples is replaced, resulting in a sliding window, which constantly updates the training set for the classifier.

C. ALDA with Error Correction (ALDEC)

The limits of unsupervised adaptive methods lie in large shifts in the data or a crossing of class distributions, corresponding to a class label switch. Unsupervised methods in principal are not capable to follow such a switch. To tackle this problem, we developed an enhancement of ALDA which is able to profit from an additional signal elicited by the presentation of wrong classification results. Error-related potentials (ERP) are widely reported in the literature: When a person's action was incorrectly interpreted, for instance by a computer, or incorrectly performed, characteristic potentials can be measured for example in EEG and fMRI [9], [10], [11].

Our Adaptive Linear Discriminant Analysis with Error Correction (ALDEC) uses a probabilistic model of a binary error signal in addition to the neuronal signal related to the action to be decoded. Based on the reliability R of the error signal, its influence on the adaptation of the decoder is weighted. This reliability could be inferred from training experiments with every subject, where one presents a random decoding error to a subject and registers, how often the corresponding error related neuronal activity is detected. In detail, ALDEC works as follows: In every step, after applying ALDA, the probabilities of ALDA are updated with the error probability. Incorporating the neuronal error signal e, equation (2) reads

$$p(c_k|x,e) = \frac{p(c_k)p(x,e|c_k)}{p(x,e)} \quad (8)$$

and assuming the neuronal signal x and the error related potential as independent

$$p(c_k|x,e) = \frac{p(c_k)}{p(x,e)} p(x|c_k)p(e|c_k). \quad (9)$$

In the ideal case where every error is detected, the distribution of an error signal is

$$p(e|c_k) = \begin{cases} 1 & \text{for } c_k \neq c_{est} \\ 0 & \text{for } c_k = c_{est} \end{cases} \quad (10)$$

depending on the real class c_k and the estimated class c_{est}. More generally, the reliability R indicates, given an error occurred, how probable the error is detected. We assume the symmetric case, where this is equal to the probability that no error was detected if no error happened. The probability that there was no classification error equals the decoding power DP. Out of this conditions the following probabilities can be deduced:

	Error occurred	No error occured
Error detected	$(1-DP)R$	$DP(1-R)$
No error detected	$(1-DP)(1-R)$	$(1-DP)R$

III. RESULTS

A. ALDA with Error signal

In simulation studies with a two dimensional signal vector, we introduced shifts of the mean values of the class distributions between training and feedback mode. First, we calculated the theoretically maximal decoding power *DPmax* by training the classifier at each trial on the actual simulated distributions. Then, we compared *DPmax* with the decoding power of LDA, ALDA and ALDEC.

Two scenarios were distinguished:

1) Class distributions rotate around the center of the vector between the two means by 90 or more degree. In this case, the error signal is needed and significantly improves the decoding power. For an example see the temporal evolution of two real $(R1, R2)$ and estimated $(E1, E2)$ class distributions in Fig.1 and the corresponding decoding power over time in Fig.2 for different classification techniques. As the real distributions in feedback are rotated 90° compared to the training, a purely unsupervised ALDA in the lowest row is not capable to follow the change. The upper row shows the ideal case of an adaptive LDA plus a 100%

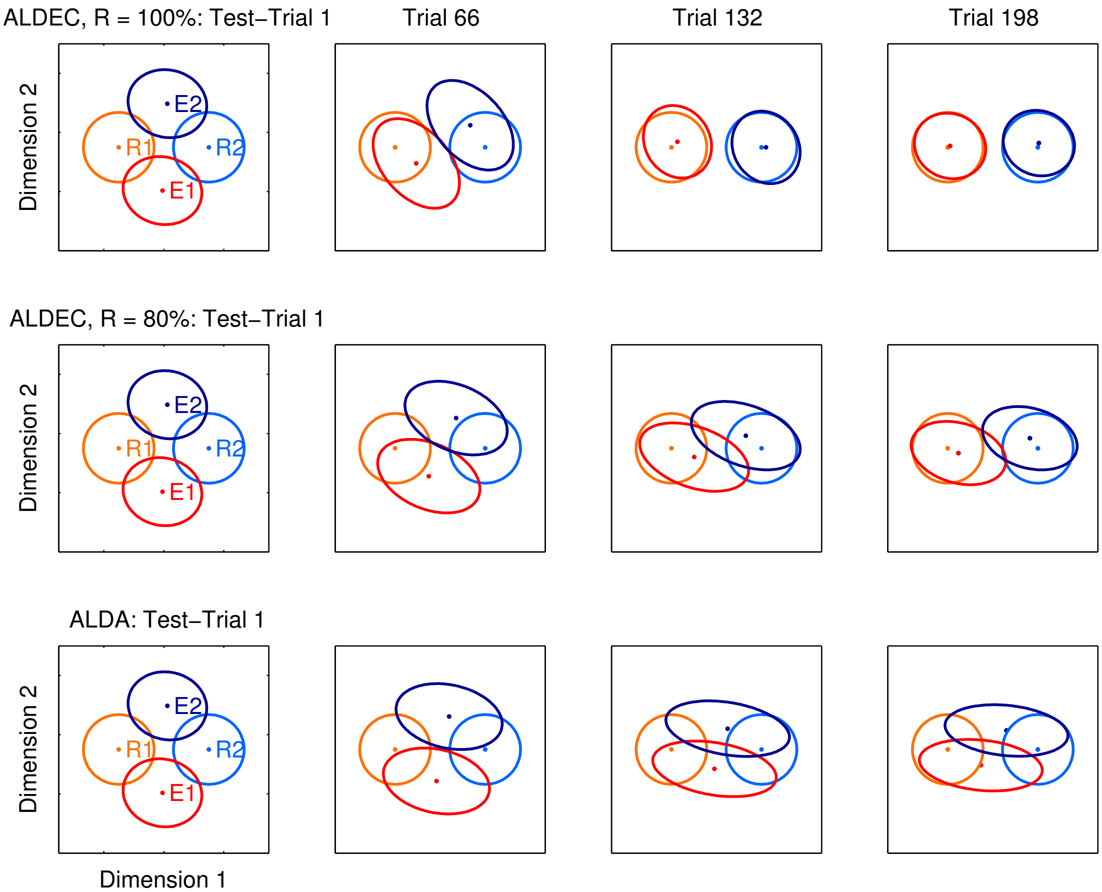

Fig. 1. **Temporal evolution of estimated class distributions for a 90° rotation between training and feedback** The left plot shows the situation at the beginning of the feedback session: The classifier is based on the training data, therefore the estimated class distributions E1 and E2 are not the same as the real feedback distributions R1 and R2. Three methods are compared: ALDEC with reliability R =100% in the first row, ALDEC with R = 80% in the second row and in the third row ALDA without the use of an error signal. In contrast to ALDEC, ALDA cannot follow the 90° rotation of two classes with mean distance 3 and covariance $\binom{1\ 0}{0\ 1}$.

reliable error signal. The estimated distributions are quickly approaching the real ones and are completely overlapping at the end. In the middle row the error signal is only 80% reliable but still has a remarkable effect. This result, which can also be seen in the corresponding decoding power values in Fig.2, indicates that even a non-perfect detection of the error will be of much help.

2) The means of the class distributions were shifted equally. In this case, ALDA already substantially outperforms LDA and after a sufficient number of trials, nearly reaches $DPmax$ as shown in Fig.3. In this case, adding an error signal does hardly increase the performance of adaptive decoding.

B. Application of ALDA to EEG-BCI experiments

To test its performance with real data, we applied ALDA to non-stationary EEG BCI data using the data set IIIb of BCI-Competition III [7]. This dataset contains information of the bipolar EEG channels C3 and C4 for cued motor imagery with online feedback and was provided by the BCI-Lab, Graz University of Technology, (G. Pfurtscheller, A. Schlögl) to

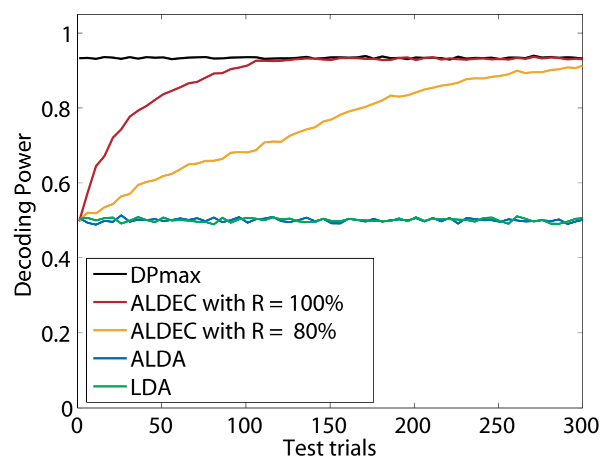

Fig. 2. **Decoding power over time for 90° rotation between training and feedback** LDA and ALDA show a decoding power at chance level 0.5. If an error signal with reliability of 100% or 80% is additionally considered, the decoding power nearly reaches DP max after 150 or 300 trials, respectively.

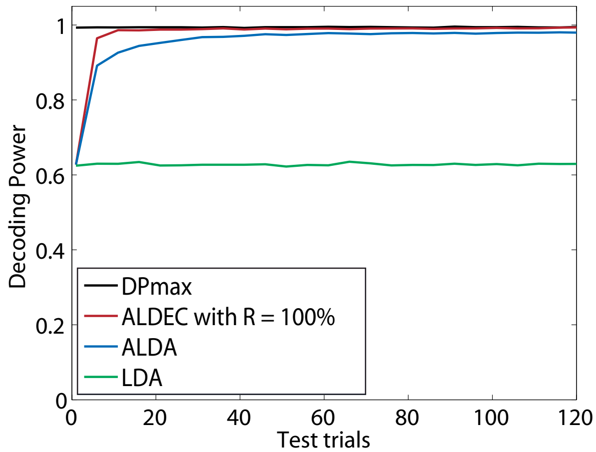

Fig. 3. **Decoding power over time for a linear shift between training and feedback**. The class distributions with distance 5 and covariance $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ were linearly shifted by 10 and rotated by 20° from training to feedback session. In this case, the performance of ALDA comes very close to the one of ALDEC 100%, and to the maximal decoding power. LDA is not able to track the changes and its decoding power stays at chance level 0.5.

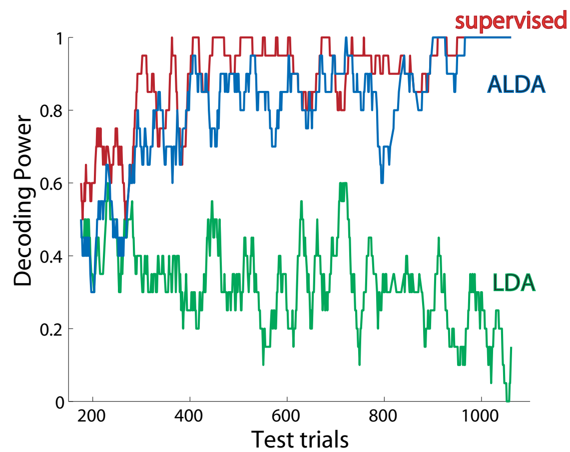

Fig. 4. **Adaptive vs. non-adaptive LDA for EEG-BCI data** In a binary task of the BCI-Competition III, data set IIIb, ALDA achieves much higher decoding power as the static LDA, nearly as high as for the supervised case.

test adaptive classifiers. We used the labels for two movement classes (left hand, right hand) of the first 200 trials to train a LDA decoder. For the following trials regarded as test data we compared three algorithms:

1) The static LDA
2) ALDA, where we updated every step based exclusively on our class label estimation (see part II.B)
3) A supervised LDA, updated with the real class labels.

Before applying the discriminant analysis, the data was smoothed with a Gaussian kernel of 400ms standard deviation and binned into a 20 dimensional vector taking data from the 2 recording electrodes every 400ms from start till end of the feedback presentation time. The decoding power is calculated by comparing the estimated labels with the real labels that were provided together with the data, averaged across a sliding window of the last 20 test trials. The result of one such analysis (subject S4b) is shown in Fig.4. For this example, ALDA (blue) clearly outperforms LDA (green) after trial 300 and the decoding power achieved by ALDA nearly reaches that of the completely supervised approach where the real labels were used (red).

IV. DISCUSSION & CONCLUSION

In a real-life BCI application it is not feasible to use supervised adaptation techniques as they always need an additional information mode to provide the real class labels. Our adaptive approach is not as accurate as the supervised one but it allows the user to act independently and it clearly outperforms the static case. Moreover, it is very flexible: while it is not necessary to know the real movement intention it is still possible to use this information, if it becomes available at certain points in time. This could be due to an additional neuronal signal that does not interfere with the patient's actions but can indicate if the decoding was correct. This might be particularly useful in cases where the shifts of the distributions are too large or complicated to be tracked by a purely unsupervised method alone. As in many relevant cases already an 80% reliable error signal significantly improves the performance, we think that the use of error related potentials to adapt the decoder is very helpful. Taken together, we demonstrated that our adaptive decoding approach for BCIs can substantially increase decoding performance in presence of non-stationary brain signals without requiring external knowledge of the subject's intentions beyond an initial training period.

REFERENCES

[1] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller and T. Vaughan: BCIs for communication and control. Clin Neurophysiol 113;767-791 (2002)
[2] B. Blankertz, G. Dornhege, M. Krauledat, G. Curio, and K.R. Müller. The non-invasive Berlin Brain-Computer Interface: Fast acquisition of effective performance in untrained subjects. NeuroImage (2007).
[3] P. Shenoy, M. Krauledat, B. Blankertz, R. Rao and K.R. Müller: Towards adaptive classification for BCI. J. Neural Eng 3, R13 (2006).
[4] D. Taylor, S. Tillery and A. Schwartz: Direct cortical control of 3D neuroprosthetic devices. Science 296, 1829 (2002).
[5] J. Carmena, M. Lebedev, R. Crist, J. O'Doherty, D. Santucci et al.: Learning to control a brain-machine interface for reaching and grasping by primates. PLoS. Biol. 1, E42 (2003).
[6] G. Gage1, K. Ludwig, K. Otto, E. Ionides and D. Kipke: Naive coadaptive cortical control, J. Neural Eng. 2, 52-63 (2005)
[7] C. Vidaurre,A. Schlögl, R. Cabeza, R. Scherer and G.Pfurtscheller: A fully on-line adaptive BCI. IEEE Trans Biomed Eng. 53(6):1214-9(2006).
[8] V. Roth and V. Steinhage: Nonlinear Discriminant Analysis Using Kernel Functions. NIPS, 568-574 (1999)
[9] V. van Veen and C. Carter: The anterior cingulate as a conflict monitor: fMRI and ERP studies. Physiol Behav. Dec;77(4-5):477-82 (2002)
[10] B. Blankertz, G. Dornhege, C. Schäfer, R. Krepki, J. Kohlmorgen et al: Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial EEG analysis. IEEE Trans. Neural Syst. Rehabil. Eng , 11, 127-131, (2003)
[11] Nieuwenhuis, Holroyd, Mol and Coles: Reinforcement-related brain potentials from medial frontal cortex. Neurosci. Biobehav. Rev (2004)