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

Continuous Detection of Motor Imagery in a Four-Class Asynchronous BCI

E. B. Sadeghian, M. H. Moradi

Abstract— **Asynchronous Brain Computer Interface (BCI) is an important class of BCI systems that has not received enough attention from the BCI community. In this work we introduce for the first time a system for classification of four different motor imageries in the context of an asynchronous BCI system which distinguishes between periods of movement imagination occurrence and idling or resting periods of ongoing EEG signal as well as classifying the 4 class motor imageries. We used two multi class extensions of the method of Common Spatial Patterns (CSP) for feature extraction and LDA, SVM, and MDA well known classifiers for combination purposes. We have applied our procedure to data set IIIa from BCI Competition III [2]. Offline evaluation of a prototype system demonstrated true positive rates in the range of 56%-88% with corresponding false positive rates in the range of 18%-9%.**

I. Introduction

BCI or Brain- Computer Interface is a communication system between human and computer unrelated to environmental nervous or muscular systems outputs. These systems suggest a new way of communication based on EEG, ECoG or MEG signals and can be divided into systems working in synchronous and asynchronous modes. The majority of BCI systems work on synchronously recorded spontaneous EEG with cue stimulus information provided. These systems called synchronous BCIs process the ongoing EEG of predefined time windows in which the imagination of movement or other mental activity has occurred and discard the signal elsewhere. Furthermore the synchronous control strategy is an unnatural and discomforting mode of interaction for most typical applications. In contrast to synchronous BCI there is what we call asynchronous BCI which is defined as a brain computer interface allowing the user to intend a specific mental activity whenever he/she wishes to. Since the ultimate goal of BCI systems is to allow the user to have complete control of her his/her external world just through imagining things and producing mental patterns in a user driven not system driven strategy, more serious advertence towards asynchronous or self-paced applications is necessary.

Manuscript received April 16, 2007.

E. B. Sadeghian is a M.Sc student with the Biomedical Engineering Department, Amir Kabir University of Technology (Tehran Polytechnic), Tehran, Iran (e-mail: e_bsadeghian@ aut.ac.ir).
M. H. Moradi is with the Biomedical Engineering Department, Amir Kabir University of Technology, Tehran, Iran (e-mail: mhmoradi@ aut.ac.ir).

An asynchronous control system should be able to discriminate the brain signal between intervals of mental activity and resting or idling periods as well as deciding between different types of supported mental activities. The main goal of this paper is to introduce the concept of multiclass paradigms in the context of an asynchronous control application and to increase the performance of such a system. Our ultimate goal is to maximize *the true positive* detections during an intended mental activity and to minimize *the false positive* detections in the idling periods as much as possible.

A two-class simulation of an asynchronous BCI has been considered in [1], which our result compares favorably with. The great care must be taken in the interpretation and evaluation of results provided here. Since in contrast to an asynchronous switch which discriminates only between intervals of having intentional control versus idling periods or having no control intention [7], [8], here we count a true positive a detection in which the true type of mental activity has been detected also. Furthermore we count a false positive whenever a non event has been detected as an event, or an event has been classified incorrectly. Therefore the evaluation criteria have become more stringent and it is important to recognize this when evaluating these results.

This paper is organized as follows. Section II describes the characteristics of dataset used in this work. In Sections II and III we briefly describe the preprocessing and CSP method of feature extraction and its extension to multi class paradigm. In section IV we introduce two classification and evaluation algorithms used in this study. Section VI through VII have our results and discuss the differences between algorithms used.

II. Dataset

We have performed our analysis on dataset IIIa from BCI Competition III [2] provided by the Laboratory of Brain-Computer Interfaces (BCI-Lab), Graz University of Technology (Prof. Gert Pfurtscheller, Alois Schlögl). This data includes cued motor imageries of 4 classes (left hand, right hand, foot, tongue) provided with 3 subjects, 60 channels and 60 trials per class. It has been filtered between 1and 50 Hz and sampled with 250 Hz. The experiment consists of several runs (at least 6) with 40 trials each. After trial begins, the first 2s were quite, at t=2s an acoustic stimulus indicated the beginning of the trial, and a cross "+" is displayed; then from second 3 to 4.25 an arrow to the left, right, up or down was displayed; at the same time the subject was asked to imagine a left hand, right hand, tongue or foot movement, respectively, until the cross disappeared at t=7s. Each of the 4 cues was displayed 10 times within each run in a randomized order. Fig. 1 depicts this paradigm.

Note that even though this dataset has been recorded synchronously with cue stimulus timing information provided, we have benefited from this information only in the final testing phase in order to validate our results, not in the proposed procedure itself.

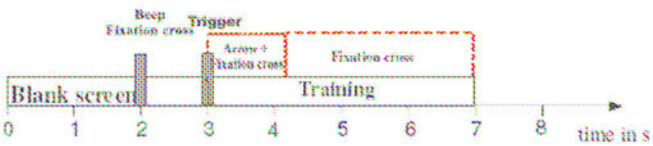

Fig. 1. Timing of the paradigm.

III. Preprocessing

In order to clean up the data from noise and artifact reduction, all 60 channels were band pass filtered from 8Hz to 30Hz applying a causal IIR filter. This broad frequency band contains the most key information used in BCI research and has shown to surpass other narrower frequency bands and improves the classification accuracy [3].

IV. Feature Extraction Methodology

A. Method of Common Spatial Patterns

The CSP method was originally proposed by Muller-gerking et al. [3] and by Ramoser et al. [4]. This method leads to a projection matrix whose rows are the discriminative spatial filters that distinct between only two conditions.

Having signals projected with projection matrix computed from training trials, the features for classification proper are vectors whose elements are the variances of the projected signals. In this method we indicate each trial with a $N \times T$ matrix, with N the number of channels and T the number of samples in time. More specifically the features we use for classification are obtained from the variances of the first and last *m* rows (*m* most discriminative features) of projected trials. Let var_p^i be the variance of the *p*-th row of *i*-th projected trial Z^i, i.e. the variance of the expansion to mode p. The feature vector for trial *i* is composed of the *2m* variances var_p^i for p running from $1 \ldots m$ and from $N-m+1 \ldots N$, normalized by the total variance or the projections retained, and log-transformed,

$$f_p^i = \log\left(\frac{\mathrm{var}_p^i}{\sum_{p=1}^{2m} \mathrm{var}_p^i}\right) \qquad (1)$$

The transformation to logarithmic values is done in order to make the distribution of elements in f^i normal.

Since the only computation we do after extraction of projection matrices are a few scalar multiplications and calculations of variances, this method is fast and suitable for online control situations. It is to be said that having these notable advantages, this algorithm has the disadvantage of requiring a large number of almost artifact free channels to produce good results.

B. CSP Multiple Extensions

As it is considered in [5], there are a few methods as CSP's extensions to multi class paradigms. In this work we have used the following two extensions:

1) *One versus the Rest extension:* In this method we compute spatial filters for each class against all others. Therefore in our 4 class problem we develop 4 projection matrices from training data. This method uses multi class classifiers to classify all projected signals.

Following this, we retain only two spatial patterns (the first and last row of captured projection matrix), which leads to an 8 dimensional feature vector.

2) *Multiple Binary extension:* This algorithm computes a projection matrix between each possible pair of classes. Therefore, for the classification of N classes $N(N-1)/2$ number of projection matrices and binary classifiers are needed.

Retaining only the most discriminative spatial filters as above provides us with a 12 dimensional feature vector for each sample to be classified.

V. Classification and Validation

We have tested our procedure with *Linear Discriminant Analysis* and *Support Vector Machine* as binary classifiers and with *Mahalanobis distance classifier* as multi class classifier.

The minimum classification error was obtained based on tenfold cross validation of all samples in 3s to 7s interval (with respect to stimulus onset) of training trials. Each classifier was trained with feature vector of sample point with minimum average classification error and has been used for continuous classification of test signal.

Following the *One versus Rest* approach of CSP feature extraction, we have used a minimum distance classifier based on *Mahalanobis* distance. Each point of continuous EEG signal is classified as belonging to the class mean to which it was closest to. Equation (4) describes the calculation of this distance from a point to a class mean.

Here, d_i is the *Mahalanobis* distance from the point x to the class mean of class i, m_i, and the inverse of covariance matrix for the class data is C_i^{-1}

$$d_i = (x - m_i)^T C_i^{-1} (x - m_i). \quad (2)$$

Having 4 classes, this classifier provides 4 distances as its outputs. Each output is responsible for the detection of the occurrence of one of the 4 movement imaginations separately e.g. if distance #1 crosses a certain threshold, the occurrence of class#1 has been detected.

As for *Multiple Binary extension* of CSP feature extraction, we have used LDA and SVM classifiers between each possible pair of 4 classes. In this approach a point is classified as a certain class if and only if all 3 classifiers involved, classify it as belonging to that class.

In our analysis each one of the motor imageries is evaluated separately. In each analysis one motor imagery, as event is compared with other imageries and idling periods as nonevent (within trial intervals or between trial intervals of having no control intention).

Therefore 4 different measures of true positive rate (TPR) and false positive rate (FPR) can be provided. These quantities are captured as following:

$$TPR = \frac{TP}{TP + FN} \quad (3)$$

$$FPR = \frac{FP}{TN + FP} \quad (4)$$

Where TP (true positive), is a true detection of movement imagination (an event); FN (false negative) is a sample belonging to an event interval that remains undetected; FP (false positive) is a false detected event or a nonevent sample that has been detected as belonging to an event interval (note that in our analysis these definitions differ from the common ones); TN (true negative) is a nonevent sample that has been classified as nonevent.

VI. Results

In order to calculate the covariance matrices of CSP method all 60 channels of EEG was used and from different window sizes, a CSP window of 1500 ms extending backward from the classification point was proved to be optimum.

In all of the results provided, the last available run of each subject has been used as testing signal while all previous runs were used for training.

A typical example of the classification accuracy captured with *Kappa* value[6] generated from training phase is shown in Fig. 2.

We have trained each classifier with feature vector related to classification point of best *Kappa* value obtained from training phase and used it for the classification of the last available run as a continuous test signal. In addition the projection matrices belonging to this point were retained for the feature extraction of the test signal.

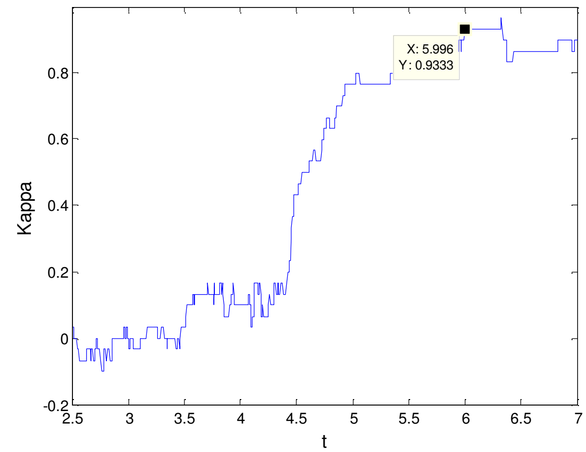

Fig. 2. Time course for classification Kappa for tenfold cross validation, plotted from 2.5s to 7s with respect to trigger point.

For each one of the motor imagery classified here the evaluation values, TP and FP rates were counted separately in samples. In each evaluation one imagery was opposed to intervals of other imageries and resting period. More specifically the time interval of 4.25s to 7s after stimulus onset of each trial was considered as an event interval while others are considered as nonevent periods of the test signal.

Results appearing here belong to subject k3b whose data has the best performance among two others. Other subject's results appear slightly under these values.

Table I shows the true and false positive rates following *One versus Rest* approach of CSP. In this algorithm only one Mahalanobis distance classifier along with 8 dimensional feature vectors has been used and considering its low cost and brief computations, compares favorably with other approaches.

TABLE I

TP and FP rates Corresponding to the Best threshold of Subject K3B and MDA Classifier In One versus Rest Approach with 8 dimensional feature vector.
The Abbreviations LH, RH, F, and T Stand for Left Hand, Right Hand, Foot, And Tongue Movement Imagery Respectively.

K3B	LH	RH	F	T
TPR%	86	83	86	54
FPR%	15	16	18	22

TABLE II

TP and FP rates Corresponding to SVM Classifier In Multiple Binary Approach with 12 dimensional feature vectors

K3B	LH	RH	F	T
TPR%	75	83	86	51
FPR%	8	9	16	24

TABLE III
TP AND FP RATES CORRESPONDING TO LDA CLASSIFIER IN MULTIPLE BINARY APPROACH
WITH 24 DIMENSIONAL FEATURE VECTORS

K3B	LH	RH	F	T
TPR%	68	84	84	53
FPR%	8	10	17	18

TABLE IV
TP AND FP RATES CORRESPONDING TO SVM CLASSIFIER IN MULTIPLE BINARY APPROACH
WITH 24 DIMENSIONAL FEATURE VECTORS

K3B	LH	RH	F	T
TPR%	88	87	83	56
FPR%	9	8	17	18

Tables II trough IV depicts the *Multiple Binary* approach of CSP, among which the SVM classifier with 24 dimensional feature vectors has the best results. In this approach we have retained the two most discriminative spatial patterns out of 60 (relating to $m = 2$ in equation (1)). In this evaluation no threshold value was used, since the positive and negative values of each binary classifier's output seemed to perform best.

VII. CONCLUSION

In this paper a novel approach for continuous classification of EEG during 4 movement imaginations in a user controlled situation was introduced.

We used two common extensions of CSP method and combined the features with different classifiers and found out that *One versus Rest* approach considering its low cost and brief amount of computation is a capable approach in online classifications.

In comparison, the multiple binary classifiers method with 6 SVM classifiers and 24 dimensional feature vectors gives the best results for all 3 subjects having the advantage of not requiring a threshold value to be determined beforehand from training data.

It should be noted that, we expected the above method to give much lower false positive rates in comparison to the other, since for detection of each class, all 3 outputs of independent classifiers count simultaneously, thus the probability of all of them being wrong at the same time is low, however, in practice it does not make a noticeable difference.

REFERENCES

[1] G. Townsend ,B. Graimann, and G. Pfurtscheller "Continuous EEG Classification During Motor Imagery—Simulation of an Asynchronous BCI" *IEEE Trans. Neural, Rehab.*vol. 12, pp. 258-265,

[2] B. Blankertz, K.-R. M¨uller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schl¨ogl, G. Pfurtscheller, J. del R. Mill´an, M. Schr¨oder, and N. Birbaumer, "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, 2006.

[3] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal spatial filters for single-trial EEG classification in a movement task," *Clin. Neurophysiol,* vol. 110, pp. 787–798, 1999.

[4] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehabil. Eng*, vol. 8, no. 4, pp. 441–446, Dec. 2000.

[5] G. Dornhege, B. Blankertz, G. Curio, and K. R. Muller, "Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 993–1002, Jun. 2004.

[6] J. Carletta, "Assessing agreement on classification tasks: the Kappa statistic," *Computational Linguistics*, vol. 22, pp. 249–254, 1996.

[7] S. G. Mason and G. E. Birch, "A brain-controlled switch for asynchronous control applications," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1297–1307, Oct. 2000.

[8] J. F. Borisoff, S. G. Mason, A. Bashashati, and G. E. Birch, "Braincomputer interface design for asynchronous control applications: Improvements to the LF-ASD asynchronous brain switch," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 985–992, Jun. 2004.