Detection of ERD/ERS in single-channel EEG with Application in Asynchronous BCI

E. B. Sadeghian, M. H. Moradi

***Abstract*— in this study we have employed most effective feature-based and model-based methods to extract motor activity changes characterized by EDR/ERS patterns in Electroencephalogram signal. We exploited for the first time Kat'z fractal dimension to identify motor related changes with application in asynchronous BCI systems. Methods applied in this study include Fractal Dimension, Band Power, m-spacing estimate of Entropy, and Quadratic Model-based detector according to a previous work in [2]. Although we assess our methods on an idle versus execution of foot movement data set, our long term goal is to employ them on imagery movement data as well. Evaluations of our early stage experiments reveals true positive rates of 76.7%-96.7% with their false positive rates of 6.45%-3.23%, respectively.**

Keywords- Asynchronous BCI, Asynchronous Brain-Controlled Switch, ERD/ERS Patterns, Fractal Dimension

I. INTRODUCTION

BRAIN-COMPUTER INTERFACE (BCI) is a system in which user's intentions are conducted towards an external device or neural prosthesis or even is used to control Functional Electrical Stimulation (FES), not requiring any physical execution. BCI systems, based on their mode of operation, are divided into two major classes, i.e. synchronous and asynchronous. Synchronous BCIs operate in a system controlled manner, where system orders the user when to start imagining executing a task. Signal processing in synchronous systems is limited within previously defined time windows, in which user is allowed to operate. On the contrary, there are asynchronous systems which allow the user to produce motor related patterns whenever he/she wishes to. Here the neurophysiological signal e.g. EEG should be continuously monitored to be able to detect "event" related patterns from "idle" thinking. In these systems, the challenging issue is to distinguish the occurrence of motor related changes from spontaneous EEG, accurately enough to have a reliable brain-controlled switch. Furthermore, the design of an asynchronous BCI has been carried out so far in two major ways. One way has been the incorporation of event detection task into classification of motor activity; by thresholding the classifier's scores [1]. Hereby functions of synchronous and asynchronous systems are gathered in one system. The veritable point is that in these systems the errors related to detection of every class of motor activities are incorporated in system's total performance. Take the example of detection of right hand versus left hand movement; here the imperfection due to detection of right hand movement occurrence versus idling accumulates with imperfection due to detection of left hand activity versus idling. Therefore system's total performance would suffer from both simultaneously.

The second way of designing an asynchronous BCI is to detect the occurrence of motor activity prior to motor classification [3]. This is the area that has not received enough attention from BCI community in contrast to motor classification which has been already well studied in literature [4], [5]. Accordingly, there are feature and model-based methods capable of detecting motor related patterns from spontaneous EEG signals. Feature-based methods employ features capable of extracting oscillatory or low frequency changes like ERD/ERS patterns or ERPs respectively. LF-ASD [3] applies a wavelet based detector to extract energy changes in 1-4Hz frequency band. But in data with no significant ERPs, as here is the case, one needs to rely on ERD/ERS extracting features as Band Power [6], Entropy, etc. In model-based methods, we fit models describing "motor related event" and "idle/rest" parts of a signal. Thereafter act of detection changes to the question of "which model fits better?"

In this paper we employ Band Power, m-spacing estimate of Entropy, Fractal Dimension from feature-based methods and "Quadratic detector" from model-based ones, to detect spectral changes of ERD/ERS patterns caused by execution of foot movement. The feature or detector signal is then fed to a threshold detector (THD) to discriminate between event and idle parts of the signal.

Manuscript received September 19, 2007.

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

II. DATASET

This data has been provided by the Laboratory of Brain Computer Interfaces (BCI-Lab), Graz University of Technology (Prof. Gert Pfurtscheller, Reinhold Scherer) [7]. Recording consists of 2 subject's EEG of 3 runs with 30 trials each. At t=0s a cross “+” was presented; then at

t=2s an arrow (pointing downwards) was displayed as cue. At the same time the subject was asked to perform foot movement. The duration was about 1s. At t=4.25s the cue and at t=6s the cross disappeared. After the trial end at t=7.5s a random inter-trial-interval of max. 1s was presented (Fig.1, left). 1-channel data is laplacian derivated over electrode position Cz (Fig.1, right). The EEG was analogy filtered between 0.5 and 30Hz and sampled with 250Hz.

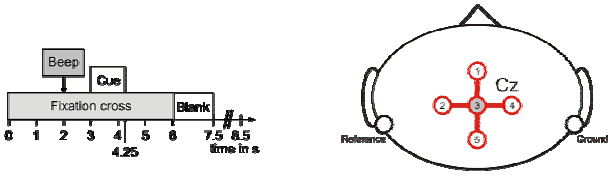


Fig. 1. Electrode positions (left) and timing scheme (right).

III. METHODS

A. ERD/ERS Visualization

Due to sensory, cognitive or motor associated events, EEG's high frequency energy fluctuates in specific regions of brain and in specific frequency bands. These spectral changes are called Event Related Desynchronization (ERD), the decrease of energy caused by correlate activation of neural system in specific cortical area and Event Related Synchronization (ERS), the increase of energy due to correlate deactivation in specific cortical area of the brain. The former is usually followed by the latter and they both are time-locked but not phase-locked with respect to event onset. ERD/ERS patterns have been widely used in motor classification [4], cognitive studies [7], and especially in BCI [2], [4], [6]. According to some researches, these patterns happen in imagery as well as real movements [8]. These patterns can be successfully visualized in time-frequency maps as images. Here we applied a bootstrap–based method [9], to extract significant ERD/ERS changes of our 1-channel data.

Energy values have been obtained by squaring samples of fast Fourier transformed data in predetermined frequency bands and then averaging over all trials. Fig.2 shows ERD/ERS maps of subjects, k6 and l1, in 0-35 Hz frequency range and entire trial's time interval. These images were computed from second run of data, which was also considered as train set. Proportional power decrease of ERD is captured in red, while power increase of ERS is captured in blue. As these figures depict, ERD starts about 2 seconds prior to foot movement onset at t=1s, whereas ERS starts almost 2 seconds after movement onset.

In addition to ERD/ERS maps, single trial band power representations are presented in Fig.3, wherein each trial's samples after band pass filtration, were squared and smoothened with a 1 second moving average FIR filter and then log transformed. The frequency bands, in which band-pass filtering of data was performed, were those who indicated strong ERD/ERS patterns in Fig.2, i.e. almost 20-35 Hz for subject k6 and 17-32 Hz for subject l1. In Fig.3 we see energy increase of ERS starting almost 2 seconds after movement onset and continuing nearly until the end of each trial.

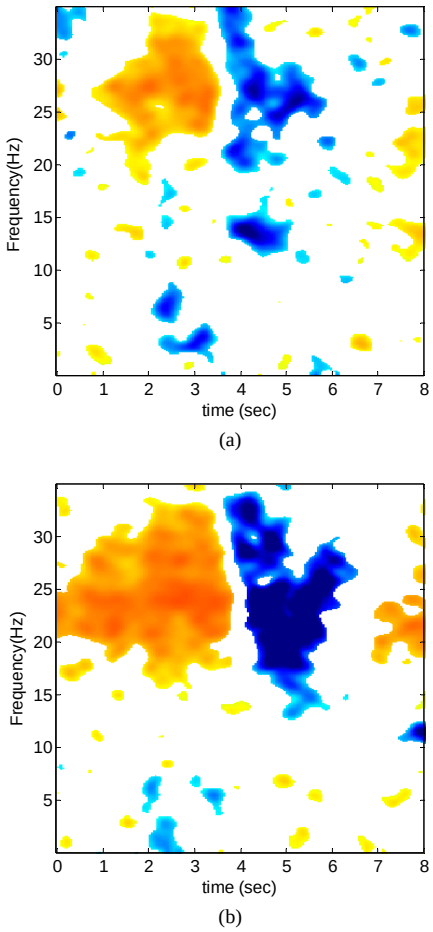


Fig. 2. ERD/ERS map of (a) subject k6 and (b) subject l1.

B. Fractal Dimension

Fractal dimension quantifies the randomness of geometric tracks in state space representation of a signal; therewith one could capture changes in oscillatory dynamics or irregularities using fractal dimension as feature. This also may be a proper candidate to depict changes of signal's energy due to desynchronization/synchronization of neurons in motor cortex.

Fractal dimension is calculated mainly by three methods among which Katz's algorithm because of its robustness

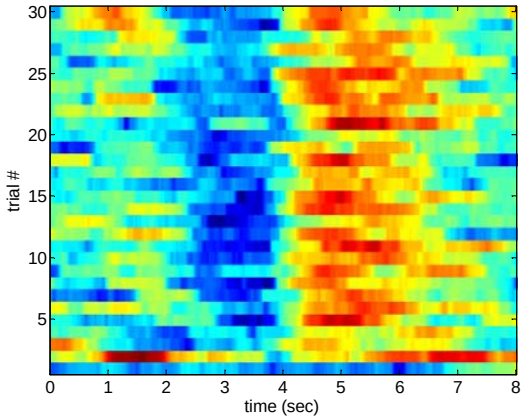
and simplicity has been considered as the most effective one [10].

Katz defines his Fractal Dimension as:

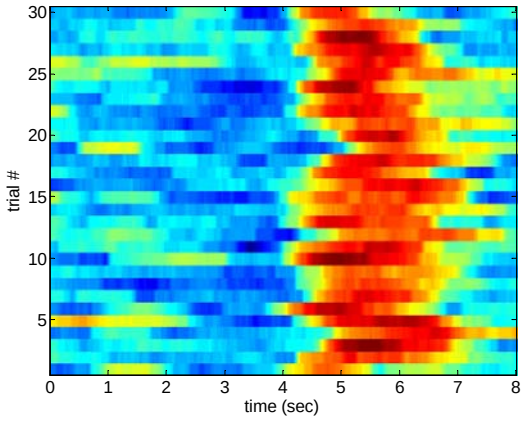
$$FD = \frac{\log(N-1)}{\log(N-1)+\log\left(\frac{d}{L}\right)} . \qquad (1)$$

Where N is the total number of data points in a data section to be analyzed; L is the total length of that specific section (sum of Euclidean distances between successive data points); and d is the diameter of the section, (the Euclidean distance between first point in the section and the point with farthest distance from).

For a signal, FD varies between 1 and 2, with 1 indicating the complete regularity or minimum energy and 2 representing maximum irregularity or energy in the context of dynamical oscillations.



(a)



(b)

Fig. 3. Single-trial band power images of (a) subject k6 and (b) subject l1 in 20-35Hz and 17-32Hz frequency bands, respectively.

C. Entropy

In signal processing, entropy is used as a measure of complexity and randomness. For a one-dimensional random variable Z, Shannon's entropy is defined as:

$$H(Z) = -\int_{-\infty}^{\infty} p(z)\log\ p(z)\,dz \cdot \qquad (2)$$

Where $p(z)$ is impulse free probability density of r.v. Z. There exist lots of estimations of entropy in literature among which we adopt an asymptotically consistent and simple to extract estimation named m-spacing entropy [11]. This estimator is based on order statistics and is computed as follows: N Samples of r.v. Z are first rearranged in non-decreasing order as: $Z^{(1)} \le Z^{(2)} \le ... \le Z^{(N)}$. A spacing of order m (m-spacing) is then defined to be $Z^{(i+m)} - Z^{(i)}$, *for* $1 \le i \le i+m \le N$. In order statistics, there is an observation in which for any random variable with an impulse-free density function $p(.)$, there exists a cumulative function $P(.)$ with uniform distribution. In other words, the "height" of cumulative curve of r.v. Z is another r.v. with uniform distribution. Accordingly, this yields:

$$E\left[P\left(Z^{(i+m)} - Z^{(i)}\right)\right] = \frac{1}{N+1}\ \forall i,\ 1 \le i \le N-1 \cdot \qquad (3)$$

This has the implication that intervals between successive order statistics have the same expected probability mass; therefore one can use this idea to develop an entropy estimator by first estimating $p(z)$ i.e. the derivative of $P(z)$ as:

$$\hat{P}(z) = \frac{1/N+1}{Z^{(i+1)} - Z^{(i)}} . \qquad (4)$$

By substituting (4) into (2) we get:

$\hat{H}(z) = \frac{1}{N-1}\sum_{i=1}^{N-1}\log\left((N+1)(Z^{(i+1)} - Z^{(i)})\right)$, which is 1-spacing estimation of entropy and suffers from high variance. This problem can asymptotically be mitigated by considering m-spacing estimate, i.e.:

$$\hat{H}(z) = \frac{1}{N-m}\sum_{i=1}^{N-m}\log\left(\left(\frac{N+1}{m}\right)(Z^{(i+m)} - Z^{(i)})\right) . \qquad (5)$$

m is typically set to $\sqrt{N}$.

D. Band Power

Signal's power in specific frequency bands is one of the most common methods to measure oscillatory activities, especially those associated with motor activity. Band power has been extracted by two major methods. One method extracts signal's power in time domain, wherein after filtration squared signal values are considered as band power features. Accordingly these values are then smoothened using a moving average FIR filter and log transformed. Second method extracts signal's power in

frequency domain. This method simply applies fast Fourier transformation to the data and stores absolute values in specific frequency band e.g. alpha and/or beta band. One recently proposed approach has incorporated phase information of signal's fast Fourier transformed version into its magnitudes [12]. This was in contrast to traditional band power features, where just magnitudes of transformation were being retained. Although it has been shown that incorporating phase information and features delayed version into traditional band power improves the accuracy of classification, here we have not followed this method because of the amount of computations and large dimensional feature vectors it introduces. Furthermore since it was not our aim to train a classifier, but rather to extract a robust decision signal, we have followed the first approach because of its simplicity and online applicability.

E. Quadratic Model

As an alternative to feature-based methods, we have employed a model-based detector following [2]. This approach assumes that each EEG segment belongs to one of the following two classes.

$$\begin{aligned} H_0 &: x \sim N(0, K_0) \quad \text{"idling"} \\ H_1 &: x \sim N(0, K_1) \quad \text{"event"} \end{aligned} \qquad (6)$$

Above models attribute normal distributions with K_0 and K_1 covariance matrices to "idle" and "event" segments of EEG signal, respectively and by assigning zero means, they ignore low frequency patterns i.e. ERPs that may exist due to motor activity. As Fig. 2 depicts, there is no significant ERP patterns in this data, therefore this model seems to be appropriate here. Accurate estimation of covariance matrices K_0 and K_1 is almost impossible using limited amount of training data, therefore quadratic detecting continues as follows. We consider a pth order autoregressive (AR) model for idle and event segments of data, separately.

$$x[n] = -\sum_{m=1}^{p} a[m]\, x[n-m] + u[n] \qquad (7)$$

Where $n > p$ is sample index. We assume $u[n]$ is independent and identically distributed (i.i.d.) white noise with zero mean and variance σ^2. Based on previous work in [2] p is set to 6.

Therefore the problem remains here is to estimate 6 AR coefficients along with noise variance σ^2 of each model separately. Here, estimation of AR parameters out of several idle/event parts of training data is slightly different. In particular we assume that all idle/event data blocks are independent from each other and consequently we construct a joint probability density function (pdf) by multiplication of all pdfs for each data block and then apply maximum likelihood (ML) estimation to this joint pdf. See Appendix for detailed explanation of AR parameters estimation. Having two separate AR models describing "idle" and "event" segments of the signal, we construct our quadratic detector as Fig. 4 stands for.

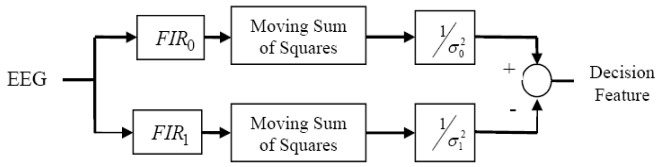


Fig. 4. Block diagram of quadratic detector.

Whitening filters FIR_0 and FIR_1 are the inverse of corresponding AR models for "idle" and "event" data blocks respectively. As EEG signal passes by these filters, it transforms to driving white noises $u_0[n]$ and $u_1[n]$ defined by:

$$u_q[n] = x[n] + \sum_{m=1}^{p} a_q[m]\, x[n-m]\ ,\ q = 0,1. \qquad (8)$$

Then a moving sum of squares computes the power of these white noises, which is then normalized by ML estimates of model variances. Whenever input signal arises from event class, the signal model of "event" fits better than "idle". So the variance of $u_1[n]$ decreases while the variance of $u_0[n]$ increases, causing a large difference captured in output decision signal.

IV. Results And Discussions

In all results provided hare, second run of data was considered as train set while third run served as test set for final evaluations. Our evaluations were executed by exerting a simple threshold to output detector signals, wherein our assessor measurements were true positive rate (TPR) and false positive rate (FPR) computed as

$$TPR = \frac{TP}{TP + FN} \qquad (9)$$

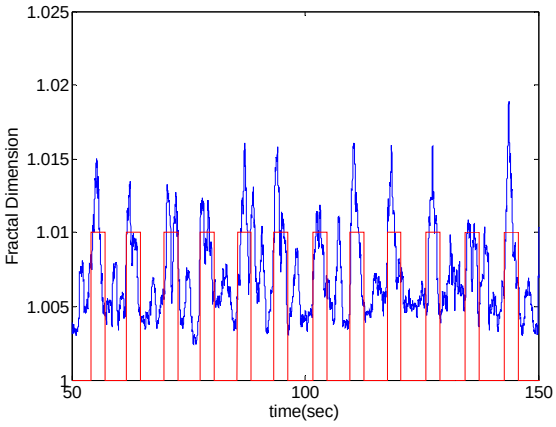
$$FPR = \frac{FP}{TN + FP}. \qquad (10)$$

Where TP (true positive), is a true detection of movement execution (an event); FN (false negative) is a sample belonging to an event interval that remains undetected; FP (false positive) is a nonevent sample that has been detected as belonging to an event interval; TN (true negative) is a nonevent sample that has been truly detected as nonevent.

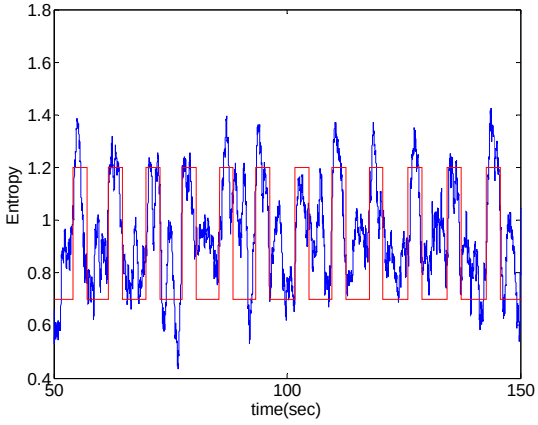
Our assessments were carried out in a sample by sample manner in which TPR and FPR were calculated using all sample points.

For feature extractions we have always applied a 1s moving window with one sample proceeding. We have empirically found the threshold and frequency band that gave best result on training data. The best threshold corresponds to the point of the ROC curve (TPR vs. FPR) closest to the line $y = 1 - x$, where it indicates equal balance between TP and FP rates.

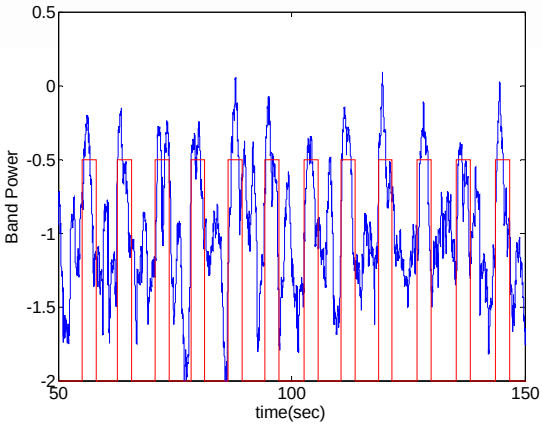
As time courses of Fig.5 depict, the increase of signal's energy happens in 4-7s interval of each trial and is captured by a rectangular pulse-shape signal. Therefore this time interval has been selected as "event" interval, wherein we count a true positive whenever detector signal exceeds the empirically predetermined threshold. A false positive is marked whenever detector signal rises above the threshold in all times excluding 4-7s interval.



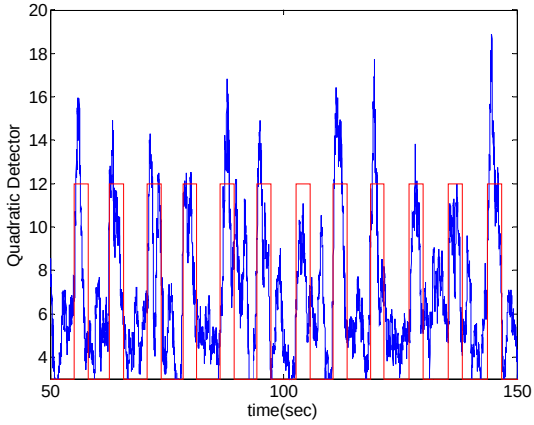
(a)



(b)



(c)



(d)

Fig. 5. Time courses of (a) Fractal dimension (b) Entropy (c) Band power (d) Quadratic model detector signals from second run of subject k6. The rectangular pulses indicate occurrence of event in 4-7s intervals.

We have summarized Area Under the ROC Curves (AUC) and TP and FP rates in Tables I, II for all methods and two subjects. The best threshold and frequency bands are selected from training set and applied on test data. As seen in the values of Tables I, II, results of Fractal Dimension is the best among others, proving the capability of this feature to work as an asynchronous switch.

TABLE I
TP AND FP RATES AND AREA UNDER THE ROC CURVE CORRESPONDING TO THIRD RUN OF SUBJECT K6

	Fractal Dimension	Entropy	Band Power	Quadratic Detector
TPR	76.7%	76.7%	73.33%	76.7%
FPR	6.45%	9.68%	3.23%	19.35%
AUC	0.933	0.929	0.933	0.876

TABLE II
TP AND FP RATES AND AREA UNDER THE ROC CURVE CORRESPONDING TO THIRD RUN OF SUBJECT L1

	Fractal Dimension	Entropy	Band Power	Quadratic Detector
TPR	96.7%	96.7%	90%	86.7%
FPR	3.23%	3.23%	3.23%	3.23%
AUC	0.986	0.985	0.985	0.984

V. CONCLUSION

It is noteworthy that although above methods are based on ERS rather than ERD and hereby introduce a time delay of several seconds into system's decision, there are BCI applications e.g. navigation in virtual environments in which not a fast response but a more accurate response is needed. Having this said, it remains our goal to compensate for this delay and to change real movement with imagery ones.

REFERENCES

[1] E. B. Sadeghian, M.H. Moradi, "Continuous Detection of Motor Imagery in a 4-Class Asynchronous BCI," *29th IEEE EMBS Annual International Conference, Lyon, France, Aug. 23-26, 2007,* to be published.

[2] Jeffrey A. Fessler, Se Young Chun, Jane E. Huggins, Simon. P. Levine, "Detection of Event-Related Spectral Changes in Electrocorticogram" in *Proc. 2th Annu. IEEE Conf. Neural Engineering*, Arlington, Virginia, pp. v-viii.

[3] 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.

[4] 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.

[5] Guido Dornhege, Benjamin Blankertz, Matthias Krauledat, Florian Losch, Gabriel Curio, and Klaus-Robert Müller, "Combined Optimization of Spatial and Temporal Filters for Improving Brain-Computer Interfacing", *IEEE Trans. Biomedical Eng*, vol. 53, no. 11,pp 2274-2281, Nov. 2006.

[6] Gert Pfurtscheller, Gernot R.Muller-Putz, Jorg Pfurtscheller, Rudiger Rupp, "EEG-Based Asynchronous BCI Controls Functional Electrical Stimulation in a Tetraplegic Patient", *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 19, pp 3152-3155, Jan. 2004.

[7] GR Müller-Putz, D Zimmermann, B Graimann, K Nestinger, G Korisek, G Pfurtscheller, "Event-related beta EEG-changes during passive and attempted foot movements in paraplegic patients.", *Brain Res*. vol. 1137, no. 1, pp 84-91, Mar. 2007.

[8] G. Pfurtscheller, C. Neuperb, C. Brunnerb, F. Lopes da Silva, "Beta rebound after different types of motor imagery in man", *Neuroscience Letters*, vol. 378, pp 156-159, Dec. 2005.

[9] B. Graimann, J. E. Huggins, S. P. Levine, and G. Pfurtscheller, "Visualization of significant ERD/ERS patterns in multichannel EEG and ECoG data," *Clin. Neurophysiol.* vol. 113, no. 1, pp. 43–47, 2002.

[10] J. Z. Liu , Q. Yang, B. Yao, R.W. Brown, G. H. Yue, " Linear correlation between fractal dimension of EEG signal and handgrip force", *biological Cybernetics*, vol. 2005, no. 93, pp. 131-140, 2005.

[11] E. G. Learned-Miller, J. W. Fisher III, "ICA Using Spacings Estimates of Entropy", *J. Machine Learning Research*, vol. 4, pp.1271-1295, 2003.

[12] George Townsend, Bernhard Graimann, Gert Pfurtscheller, "A Comparison of Common Spatial Patterns With Complex Band Power Features in a Four-Class BCI Experiment", *IEEE Trans. Biomedical Eng*, vol. 53, no. 4,pp 2274-2281, Apr. 2006.

APPENDIX. ESTIMATION OF AR PARAMETERS OF SEVERAL DATA SEGMENTS ARISING FROM IDLE/EVENT MODELS

Here we explain the mathematical part of AR parameters estimation from each model's data segments.
Consider having k independent data blocks. According to (7) and normal distribution of $u[n]$, probability density function of ith data block writes

$$P(x_i(p+1),...,x_i(N)|\, a_1,...,a_p,\sigma^2)=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{(N-p)}\prod_{n=p+1}^{N}\exp\left\{-\frac{\left(x_i(n)-a_1x_i(n-1)-...-a_px_i(n-p)\right)^2}{\sigma^2}\right\} \quad (11)$$

Where p is the model order.
According to independence of blocks

$$P_{total}=P_1P_2...P_k=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{k(N-p)}\prod_{l=1}^{k}\prod_{n=p+1}^{N}\exp\left\{-\frac{\left(x_l(n)-a_1x_l(n-1)-...-a_px_l(n-p)\right)^2}{\sigma^2}\right\} \quad (12)$$

Applying log- likelihood estimation, it reads

$$\begin{aligned}
\frac{\partial Log\ P_{total}}{\partial a_1} &= \sum_{l=1}^{k}\sum_{n=p+1}^{N}2*x_l(n-1)*\left(x_l(n)-a_1x_l(n-1)-...-a_px_l(n-p)\right)=0\\
\frac{\partial Log\ P_{total}}{\partial a_2} &= \sum_{l=1}^{k}\sum_{n=p+1}^{N}2*x_l(n-2)*\left(x_l(n)-a_1x_l(n-1)-...-a_px_l(n-p)\right)=0\\
&.\\
&.\\
&.\\
\frac{\partial Log\ P_{total}}{\partial a_p} &= \sum_{l=1}^{k}\sum_{n=p+1}^{N}2*x_l(n-p)*\left(x_l(n)-a_1x_l(n-1)-...-a_px_l(n-p)\right)=0
\end{aligned} \quad (13)$$

and,

$$\frac{\partial Log\ P_{total}}{\partial\sigma}=-\frac{k(N-p)}{\sigma}+\sum_{l=1}^{k}\sum_{n=p+1}^{N}\frac{2\left(x_l(n)-a_1x_l(n-1)-...-a_px_l(n-p)\right)^2}{\sigma^3}=0$$

Solving system of equations (13) yields to AR parameters.