Motor Imaginary Signal Classification Using Adaptive Recursive Bandpass Filter and Adaptive Autoregressive Models for Brain Machine Interface Designs

Vickneswaran Jeyabalan, Andrews Samraj, and Loo Chu Kiong

Abstract—The noteworthy point in the advancement of Brain Machine Interface (BMI) research is the ability to accurately extract features of the brain signals and to classify them into targeted control action with the easiest procedures since the expected beneficiaries are of disabled. In this paper, a new feature extraction method using the combination of adaptive band pass filters and adaptive autoregressive (AAR) modelling is proposed and applied to the classification of right and left motor imagery signals extracted from the brain. The introduction of the adaptive bandpass filter improves the characterization process of the autocorrelation functions of the AAR models, as it enhances and strengthens the EEG signal, which is noisy and stochastic in nature. The experimental results on the Graz BCI data set have shown that by implementing the proposed feature extraction method, a LDA and SVM classifier outperforms other AAR approaches of the BCI 2003 competition in terms of the mutual information, the competition criterion, or misclassification rate.

Keywords—Adaptive autoregressive, Adaptive bandpass filter Brain Machine Interface, EEG , Motor imaginary

I. INTRODUCTION

BRAIN Machine Interfaces (BMIs) or Brain Computer Interfaces (BCIs) are devices designed with the intention of helping people who are disabled in nature. This device enables a disabled person to communicate with a computer or machine using their brains' electrical activities as the only medium.

Various techniques are available for capturing brain activities, which includes electroencephalogram (EEG), functional Magnetic Resonance Imaging (fMRI), Magneto-encephalography (MEG), Position Emission Tomography (PET). Among these techniques, EEG is the most preferred for BCI designs, because of its non-invasiveness, cost

Vickneswaran Jeyabalan , is with the Faculty of Engineering and Technology, Multimedia University, Melaka,Malaysia (corresponding author phone: 60122557238; e-mail: vickneswaran.jeyab07@mmu.edu.my).

Andrews Samraj, is a lecturer with the Faculty of Information Science and Technology, Multimedia University, Melaka, Malaysia. (e-mail: Andrews.samraj@mmu.edu.my).

Loo.CK, is the Dean of the Faculty of Information Science and Technology, Multimedia University, Melaka, Malaysia. (e-mail: ckloo@mmu.edu.my).

effectiveness, and easy implementation [1, 2, 3]. This was among the reasons which encouraged us to concentrate in the design of an EEG-based BCI for this research.

A BCI's design is usually realized by using visual evoked potentials (VEP) or movement related potentials (MRP). The work presented in this paper concentrates in the usage of MRPs for the EEG-based BCI. Clear representations of MRP can be observed in the EEG's mu-rhythm (8-12Hz) and/or Beta rhythm when a person performs a motor activity or imagines a motor activity [4]. Such an activity can be easily captured from the EEG's channel C3, and C4 [5].

Current designs of BCI usually consists of four main stages, which are, raw signal acquisition, signal pre-processing or conditioning, feature extraction and finally classification of features into intended actions. Among these stages, feature extraction and classification methods plays an important role because, a successful BCI depends on its ability to extract EEG features according to different tasks and to efficiently classify them in a real time environment [6].

One of the most challenging tasks in designing a BCI is in choosing the relevant features from the EEG signals, which are chaotic in nature. Accurate feature extraction method is very important in determining the performance of a classifier. Incorrect features could lead the classifier to have poor generalization, computational complexity and requires a large number of training data set to achieve a given accuracy [7]. So, the goal of this paper is to develop and efficient feature extraction methodology.

In an EEG-based BCI, time frequency analysis methods are usually employed. Feature recognition in this paradigm, could be realized by implementing estimation procedures, which are categorized as parametric and non-parametric. Parametric approaches such as the Autoregressive (AR) and its variant Adaptive Autoregressive (AAR) modeling are the most commonly used [8,9] in the analysis of EEG signals. Parametric approaches' using AR model is well studied and is attractive because of its capability to summarize information concisely and translates them into feature vectors. Besides that, AR parameters are also mainly used due to its ability to describe the stochastic nature of EEG and it does not require any *priori* information of any relevant frequencies. Estimation

of AR parameters can be done using a wide range of stationary process such as the Burg's and Yule-Walker algorithms, [10] but EEG signals are well known for its non-stationary behavior. So, a moving window is usually employed on the signal to estimate its AR parameters. This method becomes computationally intensive, depending on the size and shift of the window. In order to overcome this problem, the kalman filtering method [11] is used instead of moving windows to estimate the non-stationary AR parameters. Kalman filter is an efficient recursive filter that estimates the state of a dynamic structure from a non-stationary signal like the EEG. The Kalman filter requires only information of the prior state of the signal and thus becomes computationally efficient. The combination of Kalman filtering in estimating AR parameters produces the methodology of Adaptive Autoregressive (AAR) modeling.

However, in reality EEG are generated by a non-linear system which consists of post synaptic neurons firing action potentials,[6] that make the signal very noisy and chaotic. Thus making the signal to have many sinusoidal components of different features and through nonlinear interactions, the signal produces one or more sinusoidal components at sum and difference frequency [12]. This makes the characterization process of features by the AR/AAR models completely difficult and could lead to very bad classification performances.

To overcome this problem, this paper introduces a new method of feature extraction, which incorporates AAR parameters and Adaptive Recursive Bandpass Filter (ARBF) to produce better classification results. Figure 1 depicts the BCI design using the proposed method. The ARBF is designed to trail the centre frequency of the dominant EEG signal and requires only one coefficient to be updated [13,14]. This is done in order to adjust the centre frequency of the filter band pass and to be approximated with the input signal [13, 14]. The time function of the coefficient then represents the dominant EEG signal. These signals are then introduced to the AAR models to estimate its parameters which will be used for classification. The ARBF method makes the characterization process of the autocorrelation functions in the AAR models better, as it enhances the EEG signal.

To evaluate the effectiveness of the proposed method, the support vector machine (SVM) and linear discrimination analysis (LDA) was used to classify the Graz BCI dataset which was used in the BCI Competition 2003 [15]. The objective of the competition is to classify imagined left or right hand performed at a particular time or trial. The performance of the classifiers are usually determined by its error rate, however, it only considers the sign of the classifier output but not the degrees of memberships of patterns belonging to each class. Therefore, error rate just provides the classification accuracies of the used classifier but not the information of how much confidence about the classification result is [6]. So, in order to overcome this limitation, entropy based mutual information (MI) was used as the evaluation criterion on the Graz data set for the BCI competition 2003

[10, 16]. The MI enables us to view the combination of classification accuracy and confidents of the classifier's output. Higher MI of the classification result indicates that the classifier produces results with higher confidence. In this paper, the classification results obtained from the proposed method is compared with results from the BCI competition 2003.
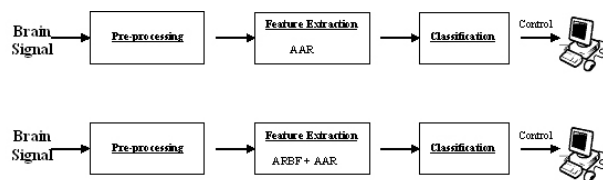

Fig. 1 (Top) BMI design using AAR for feature extraction.
(Bottom) BMI design using proposed method for feature extraction

II. FEATURE EXTRACTION

A. Adaptive Recursive Bandpass Filter (ARBF)

The adaptive recursive bandpass filtering as proposed by Gharieb is employed to estimate and track the centre frequency of the dominant signal of each EEG channel [13]. In this paper we concentrate in detecting the mu-rhythm (8-12Hz) since it is the dominant signal from the EEG channels C3, and C4. The ARBF updates only one coefficient in order to adjust the centre frequency of the filter band pass to be matched with the input signal. The time function of the coefficient represents the dominant frequency of the signal, which is feasible to be used as input for the AAR model.

B. Filter Structure

A fourth order Butterworth band pass filter is employed as the adaptive filter. The filter function, T (z) could be expressed as [17,19].

$$T(z) =$$

$$\frac{D_0 + D_2 Z^{-2} + D_4 Z^{-4}}{1 + F_1 C(n) Z^{-1} + (F_2 C^2(n) + F'_2) Z^{-2} + F_3 C(n) Z^{-3} + F_4 Z^{-4}} \quad (1)$$

where

$$D_0 = D_4 = 1/(l^2 + \sqrt{2}l + 1), \quad D_2 = -2D_0,$$

$$D_1 = D_3 = -4D_0, \quad F_1 = -2l(2l + \sqrt{2})D_0,$$

$$F_2 = 4l^2 D_0, \quad F'_2 = 2(l^2 - 1)D_0,$$

$$F_3 = 2l(-2l + \sqrt{2})D_0, \quad F_4 = (l^2 - \sqrt{2}l + 1)D_0,$$

$$l = \cot an(\pi BP)$$

The coefficient C (n) could be expressed as

$$C(n) = \frac{\cos(\pi(H_2(n) + H_1(n)))}{\cos(\pi BP)} \qquad (2)$$

where

$H_1(n)$ = normalised low cut off frequency,

$H_2(n)$ = normalised high cut off frequency,

BP = normalised bandwidth of the filter

BP is assumed to be a constant value. Based on equation 1 and 2, it can be seen that $C(n)$ is the only coefficient that has to be updated by the adaptive filter since it is also the only coefficient which is dependant with the centre frequency, $(H_1(n) + H_2(n))/2$. Hence the filter has only one centre frequency dependent, $C(n)$ to be updated.

C. Adaptive Algorithm

The output power of the filter is maximised in order for the filter, $T(z)$ to be self-adjusted to the centre frequency of the input signal, the adaptive filter coefficient, $C(n)$ is updated for the maximization of the expected output power $E\{y^2(n)\}$ [13,14]. This step can be applied by implementing a standard gradient approach [13, 14]. An algorithm called recursive maximum mean-squared (RMXMS) is used to update the filter coefficient [13,14]. The RMXMS algorithm can be described as:

$$C(n+1) = C(n) + 0.5\mu_n \nabla(E\{y^2(n)\}) \qquad (3)$$

where,

$\mu_n > 0$ = Normalized step size.

$\nabla(E\{y^2(n)\})$ = True gradient respective to $C(n)$.

Using the instantaneous gradient $\nabla\{y^2(n)\}$ as a stochastic approximate for the true gradient $\nabla(E\{y^2(n)\})$, the following equation is obtained:

$$C(n+1) = C(n) + \mu_n y(n)\alpha(n) \qquad (4)$$

where the $\alpha(n) = \nabla y(n)$ and the filter output $y(n)$ is given by:

$$y(n) = -D_0 x(n) + D_2 x(n-2) + D_4 x(n-4) - F_1 C(n) y(n-1)$$
$$-(F_2 C^2(n) + F'_2) y(n-2) - F_3 C(n) y(n-3)$$
$$-F_4 y(n-4) \qquad (5)$$

and $\alpha(n) = \dfrac{\partial y(n)}{\partial C(n)}$ can be computed from the equation above, which yields:

$$\alpha(n) = -F_1 y(n-1) - 2F_2 C(n) y(n-2) - F_3 y(n-3)$$
$$-F_1 C(n)\alpha(n-1) - (F'_2 + F_2 C^2(n)\alpha(n-2))$$
$$-F_3 C(n)\alpha(n-3) - F_4 \alpha(n-4) \qquad (6)$$

The normalized step-size μ_n is given by $\mu_n = \mu / p(n)$

where μ is a fixed positive step size. $p(n)$ is a recursive estimate of the power of the gradient with the equation:

$$p(n) = \lambda p(n-1) + \alpha^2(n) \qquad (7)$$

Where $0 << \lambda < 1$ is a forgetting factor and finally it proves that the stability is guaranteed if the update $|C(n+1)| \geq 1$, then $C(n+1) = C(n)$.

The adaptive filter further enhances the feature and provides good on-line information of the feature's distinct behaviour. The adaptive filter becomes unstable for some low frequency waves. This is because the adaptive filter is a band pass filter and the EEG signal is a low pass signal [13,14]. So, in order to solve this problem a high frequency shifting process is employed to shift the EEG frequency to the highest ones before the adaptive filtering [13,14]. The time function, $C(t)$ of the updated co efficient $C(n)$ is then introduced to the AAR model to extract features.
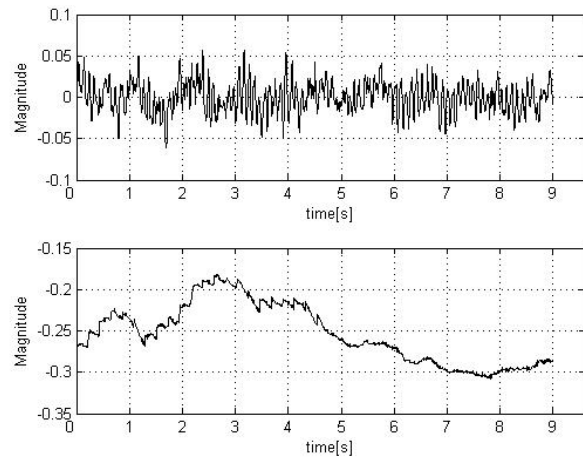

Fig. 2. Average of 10 trial realisations of raw motor imaginary signals from channel C3 (Top). Mean values of the extracted features after performing ARBF (bottom).
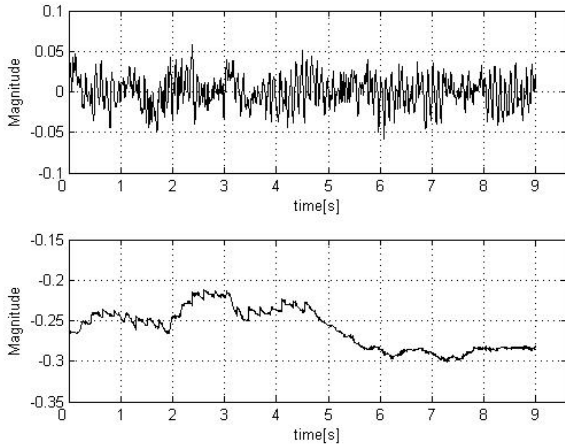
Fig. 3. Average of 10 trial realisations of raw motor imaginary signals from channel C4 (Top). Mean values of the extracted features after performing ARBF (bottom).

D. AAR Parameters

An AAR model with n^{th} order can describe the time function, C(t), of the coefficient which represents the dominant frequency of the signal in the following form

$$x(i) = y(i)^t * z(i-1) + w_j(i) \tag{8}$$

i = discrete time index
x = C(t),EEG signal after passing through the ARBF
y = n^{th} element column vector of AR coefficient
z = last element vector of nth EEG samples
w_j = zero mean, white noise process

$$y(i) = y(i-1) + w_n(i) \tag{9}$$

w_n = zero mean white noise

From equation 1 we could observe that the AR co-efficient changes with time, t in order to capture the non-static behavior of the EEG signal. Equation 1 can also be considered as the measurement model of the kalman filter, [19] while equation 2 describes the dynamics of the kalman model and we could also observe that it is modeled as random walk with an assumption of small changes in the state. The AAR estimation using Kalman filtering algorithm is shown in equation 3 to 7.

$$A(i) = (1 - uc) * A(i-1) \tag{10}$$
$$B(i) = M(i-1).z(i) / A(i) \tag{11}$$
$$\hat{y}(i+1) = \hat{y}(i) + B(i)\upsilon(i) \tag{12}$$
$$N(i) = V * trace(M(i-1)) / n \tag{13}$$
$$M(i) = M(n-1) - B(i)^t.z(i-1).M(i-1) + N(i) \tag{14}$$

$A(i)$ = Innovation process variation
$B(i)$ = kalman gain
$M(i)$ = Predicted error matrix
$z(i)$ = measurement matrix of last, n, samples of signal, x.
$y(i)$ = Estimates of AAR parameter
$N(i)$ = process noise variance
V = Updated coefficient
$\upsilon(i)$ = innovation process

The best estimation of AAR parameters for EEG signal, x, depends on the updated coefficient, V, and model order ,n. These parameters could be obtained by minimizing the variance of the prediction error. The relative error variance, R, [20] was used in this case.

$$R = (\frac{\sum \upsilon(i)^2}{S}) / \mathrm{var}\,iance(x) \tag{15}$$

Equation 8 demonstrates the mean squared error normalized by the signal variance, where S is the number of samples in the trial. This results a feature vector $D_t^{(i)}$ with a dimension of $2*n$ for every trial i and every sampling point t.

$$D_t^{(i)} = [\hat{y}_{(i),C3}^{(i),C4}] \tag{16}$$

III. CLASSIFICATION

LDA and SVM linear classifiers are used to classify the imagined hand movements in order to show the performance of the proposed method.

A. LDA Classifier

The feature vectors $D_t^{(i)}$ are first mapped into a linear transformation, where a weight vector w_t and offset w_0 were found with which the distance d was computed as the following equation:

$$d_t^{(i)} = w_t^T \cdot D_t^{(i)} - w_0 \tag{18}$$

The w_t and w_0 were obtained by maximizing the ratio of between-class variance and within-class variance to ensure the maximal distinction [21]. The within-class variance could be equated as :

$$J_w = \sum_{i=1}^{K} \sum_{l=1}^{L_i} (D^l - \mu_i)(D^l - \mu_i)^T \tag{19}$$

314

Where,

K = number of classes

μ_i =mean of class i

L_i =number of samples within class i

Whereas, the between-class variance could be equated as :

$$J_b = \sum_{i=1}^{K} (\mu_i - \mu)(\mu_i - \mu)^T \tag{20}$$

Where,

μ =mean of the training samples set.

The obtained distance $d_t^{(i)} > 0$ and $d_t^{(i)} < 0$ shows that the trial i is classified as imagined left and right trial respectively.

B. SVM Classifier

The aim of a SVM classifier is to use a discriminant hyper plane in order identify different classes [22, 23]. SVM maximizes the margins of the hyper plane (distance from the nearest training point) in order to increase the generalization capabilities [22, 23]. A regularization parameter is also used by SVM that enables accommodation to outlier and allows error on the training set [24]. This type of SVM enables classification using linear decision boundaries and also known as linear SVM. Non-linear decision boundaries can also be created by using the 'kernel trick', where it consists of implicitly mapping the data into another space by using a kernel function. The output of a binary SVM can be equated as follows:

$$Q = \mathrm{sgn}(\sum_{i=1}^{N} \alpha_i Q_i K(x_i, x) + b) \tag{21}$$

Where,

$\{x_i, Q_i\}_{i=1}^{N}$ = training samples

$x_i \in R^d$ = input vector

$Q_i \in \{-1, 1\}$ =class labels

$\alpha_i \geq 0$ = Langarian multipliers, obtained from quadratic optimization problem

b = bias

$k(x_i, x_j)$ = kernel function of SVM

The most commonly used kernel function in BCI classifications is the Gaussian or radial basis function (RBF), which could be expressed as below:

$$K(x_i, x_j) = \exp\left(\frac{-\left\| x_i - x_j \right\|^2}{2\sigma^2} \right) \tag{22}$$

SVM has several advantages in terms of good generalization properties [23,25] because of the margin maximization and regularization and insensitive to over training [25].

IV. EXPERIMENTAL RESULTS

A. Data Set

The proposed method was trialed on the BCI Competition 2003, dataset IIIb [15]. The data set was provided by the Department of Medical Informatics, Institute for Biomedical Engineering, University of Graz. The signals were obtained from a 25-year-old female relaxing on a chair with armrest. The task was to control a feedback bar by means of imagining left hand of right hand movements. The data was acquired from the EEG channels C3, Cz, and C4 (Figure.5), which was band pass filtered for a frequency range of 0.5 to 30Hz and sampled at 128 Hz. The experiment consists of 7 runs with 40 trials each. All runs were conducted on the same day with several minutes break in between. The data has a total of 280 trials, which consists of 140 labeled and 140 unlabeled trials with an equal number of left hand and right hand movements. Each trial consists of duration of 9 seconds. At the 3rd second a visual cue (Figure.4), an arrow pointing left or right is presented to indicate left or right motor movements is to be imagined.
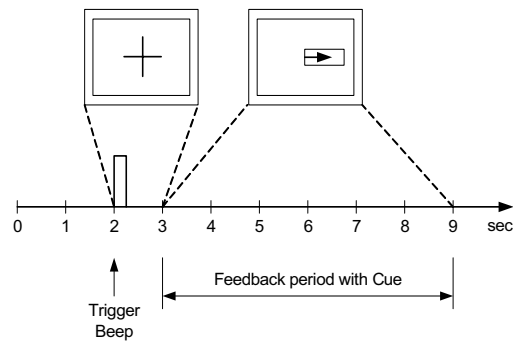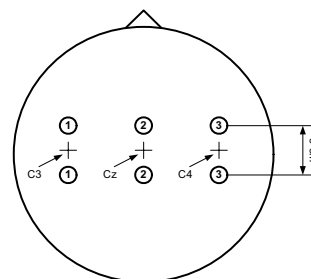

Fig. 4. The timing scheme

Fig. 5. Electrode positions on the scalp

B. Signal Analysis

In order to begin the signal analysis, the EEG signals of channel C3 and C4 from the data set are extracted. Signals from channel Cz is ignored because it contains very little significant discriminative features [26]. Signals from EEG channel C3 and C4 are first pre-processed by using a 5th order low pass Butterworth filter where the pass band is 12Hz with less than 1 dB of ripple with at least 6 dB of attenuation.

Then the pre-processed signal is introduced to the ARBF in order to track the dominant wave existing in the signal. The observed signal is composed of mu-rhythm (8-12Hz) for a period of 9 seconds for each trial. Based on the definition of section 2.1, the normalized carrier frequency is set to 0.30, while the forgetting factor and step size are set to 0.9 and 0.95 respectively. The normalized frequency bandwidth, *B*, is set to 0.15. The initial values of *r(0)* and *w(0)* is set to 100 and 0.0 respectively. Fig.2 and Fig.3 shows the results of the ARBF for the training data set. Investigating the results we could see that the time frequency of the adaptive coefficient, *w(t)*, gives clear representation of the EEG signals.

The signals are then used to estimate the AAR parameters. A model order, *n*=5 and update coefficient, *V*=0.0055 are used. The AAR parameters were estimated for every sample time point for the EEG channels C3 and C4 respectively. This resulted in a feature vector with a dimension of 2*5 for every trial and every sampling point. The features were extracted for all the 280 trials of 9 seconds each, in which 140 is the training data set and 140 is test data set.

In this paper, we used the Gaussian kernel based SVM [27]. In order for SVM to find the optimal Gaussian parameter, σ and the regularization parameter C, the genetic algorithm (GA) was used to select the suitable hyper-parameters [6]. The MI [10, 16] was used as a criterion to evaluate the performance of the proposed method with the results of BCI competition 2003's AAR parameters method.

As mentioned in section 3.1, the data set used for our experiments consists of 140 training and 140 test dataset. In order to further evaluate the performance of our proposed method, we first tested on the 140 training dataset. A 5 fold cross validation approach was employed for this procedure. The data set (140 trials) were split into five parts with equal number of trials. This procedure is repeated for five times, where for each time five different parts were trained and the remaining part was used for testing. This leave one out method is done in order to increase the validity of the classification results. Next, we used another Graz BCI motor imaginary dataset [28], to validate the proposed method. The data set was generated with different experimental settings from 3 subjects. Further information of the data set can be obtained from [28]. In our experiments, we used the data set from only one subject, O3 and 120 sample trials we chosen randomly. A 5 fold cross validation approach was also employed similarly to the previous experiment.

Finally, in order to test the proposed method on the BCI

Completion data set, we used the 140 training data set to train the classifiers and then we used the remaining 140 test data set to test the classifiers.

TABLE I
EXPERIMENTAL RESULT OF THE PROPOSED METHOD ON THE 140 TRAINING TRIALS USING 5 FOLD CROSS VALIDATION APPROACH

Feature Extraction	AAR			ARBF+AAR		
Classification	Error (%)	MI	SNR	Error (%)	MI	SNR
SVM	17.1	0.27	0.45	12.1	0.51	1.04
LDA	18.6	0.27	0.46	11.4	0.53	1.09

TABLE II
EXPERIMENTAL RESULT OF THE PROPOSED METHOD ON THE O3 DATA SET USING 5 FOLD CROSS VALIDATION APPROACH

Feature Extraction	AAR			ARBF+AAR		
Classification	Error (%)	MI	SNR	Error (%)	MI	SNR
SVM	22.5	0.24	0.39	10.8	0.46	0.88
LDA	24.2	0.25	0.41	13.3	0.44	0.85

TABLE III
EXPERIMENTAL RESULT OF THE PROPOSED METHOD AND THE BCI COMPETITION 2003 METHODS USING THE AAR PARADIGM .

Methods	MI	Error (%)	SNR
ARBF+AAR+LDA	0.49	14.10	0.97
ARBF+AAR+SVM	0.46	14.30	0.90
BCI_Competition 2003 (GroupF) AAR+LDA	0.46	15.71	0.90
BCI_Competition 2003 (GroupB) AAR+LDA	0.45	17.14	0.86
BCI_Competition 2003 (GroupG) AAR+ NEURAL NET	0.29	17.14	0.50

V. CONCLUSION

In this paper a new feature extraction method was introduced in order to classify EEG signals according to left or right hand motor imaginary. The feature extraction method includes the combination of ARBF and AAR models to produced better features for classification. Experimental results have shown that the proposed method, which is by introducing the ARBF before estimation of AAR parameters increases the performances of the autoregressive characterization function in AR models. This leads to a better classification results, especially with the usage of LDA and SVM as classifiers. Experimental results have also shown that the proposed methods are suitable to used for an efficient Brain Machine Interface design.
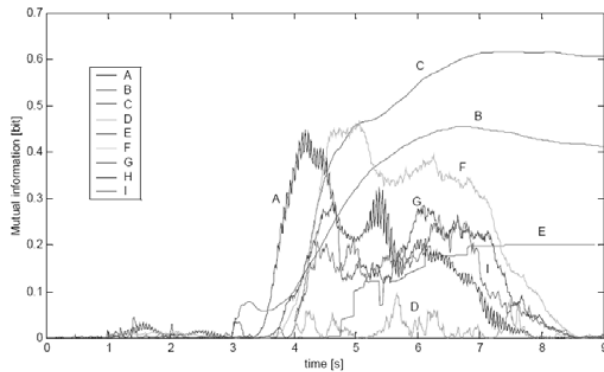
ACKNOWLEDGMENT

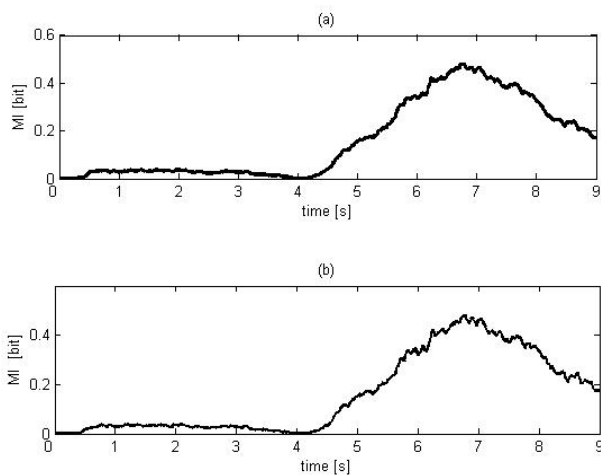Fig. 6. MI of the BCI Competition 2003 results.

Fig. 7. (a) MI for proposed method with LDA as classifier, (b) MI for proposed method with SVM as classifier.

REFERENCES

[1] L. Kauhanen, T. Nykopp, J. Lehtonen, P. Jyla¨nki, J. Heikkonen, P. Rantanen, H. Alaranta, M. Sams, "EEG and MEG brain computer interface for tetraplegic patients", *IEEE Trans. Neural Syst. Rehabil. Eng.* 14 (2) (2006) 190–193.

[2] J. Kronegg, G. Chanel, S. Voloshynovskiy, T. Pun, "EEG-based synchronized brain–computer interfaces: a model for optimizing the number of mental tasks", *IEEE Trans. Neural Syst. Rehabil. Eng.* 15 (1) (2007) 50–58.

[3] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, "Brain–computer interfaces for communication and control", *Clin. Neurophysiol.* 113 (2002) 767–791.

[4] G. Pfurtscheller, "EEG event-related desynchronization (ERD) and event-related synchronization (ERS), Electroencephalogr.: Basic Prin", *Clin. Appl.* (1999) 958–967.

[5] J. Vickneswaran, S. Andrew, CK. Loo, " Classification of Motor Imaginary Tasks Using Adaptive Recursive Bandpass Filter", *Proc. International Conference on Signal Processing and Multimedia Applications(SIGMAP2008)*,(2008) 113-118.

[6] SM Zhou, John Q. Gan b, F. Sepulveda, "Classifying mental tasks based on features of higher-order statistics from EEG signals in brain computer interface", *Inform. Sci.* 178 (2008) 1629–1640

[7] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artif. Intell.*, vol. 97, pp. 245–271, Dec. 1997.

[8] B. H. Jansen, J. R. Bourne, J. W. Ward, "Autoregressive estimation of short segment spectra for computerized EEG analysis," *IEEE Trans.Biomed. Eng.*, 28(9), 1981.

[9] G. Florian and G. Pfurtscheller, "Dynamic spectral analysis of event related EEG data," *Electroenceph. Clin. Neurophysiol.* 95, (1995) 303-309.

[10] A. Schlogl, C. Neuper, G. Pfurtscheller, "Estimating the Mutual Information of an EEG-based Brain-Computer Interface," *Biomedizinische Technik,* 47, (2002) 3-8.

[11] M. Arnold, W. Miltner, H. Witte, R. Bauer, and C. Braun, "Adaptive AR Modeling of Non-stationary Time Series by Means of Kalman Filtering", *IEEE Trans. Biomed. Eng.,* 45(5), 1998.

[12] J. Muthuswamyand, R.J. Roy, "The use of fuzzy integrals and bispectral analysis of the electroencephalogram to predict movement under anesthesia", *IEEE Trans. Biomed. Eng.* 46 (3) (1999) 291–299.

[13] R. R. Gharieb and A. Cichocki., "Segmentation and tracking of the electro-encephalogram signal using an adaptive recursive bandpass filter", *Medical & Biological Engineering & Computing 2001,39,* (2001) 237-248.

[14] R. R. Gharieb and A. Cichocki, "On-line EEG classification and sleep spindles detection using an adaptive recursive bandpass filter", *In proceedings Acoustics, Speech, and Signal Processing (ICASSP'01) IEEE International Conference,* Vol 2, (2001).

[15] BCI Competition 2003. <http://ida.first.fraunhofer.de/projects/bci/competition/>.

[16] A. Schlogl, C. Keinrath, R. Scherer, G. Pfurtscheller, "Information transfer of an EEG-based brain–computer interface", *Proc. International IEEE EMBS Conference on Neural Engineering,* (2003) 641–644.

[17] R.V. Raja Kumar and R.N. Pal, "A gradient algorithm for center-frequency adaptive recursive band pass filters", *Proc. IEEE,*73,(1985) 371-372.

[18] R.V. Raja Kumar and R.N. Pal., "Tracking of bandpass signals using center-frequency adaptive filters", *IEEE Trans. Acoust. Speech signal processing.*, 38, (1990) 1710-1721.

[19] S. Haykin, *Adaptive Filter Theory*, 4th edition, Upper Saddle River, NJ: Prentice Hall, 2002, Ch 10.

[20] A. Schlogl, S. Roberts, G. Pfurtscheller, "A criterion for adaptive autoregressive models," *Proceedings of the 22nd EMBS International Conference,* (2000)1581-1582.

[21] R.A. Fisher, "The use of multiple measurements in taxonomic problems", *Ann. Eugen.* 7 (1936) 179–188.

[22] Burges C J C , "A tutorial on support vector machines for pattern recognitionition", *Knowl. Discov. Data Min.* 2 , (1998) 121–67

[23] Bennett K P and Campbell C, " Support vector machines: hype or hallelujah?", *ACM SIGKDD Explor. Newslett.* 2 ,(2000) 1–13.

[24] F Lotte, M Congedo, A L´ecuyer, F Lamarche , B Arnaldi1, "A review of classification algorithms for EEG-based brain–computer interfaces", *J. Neural Eng.* 4 (2007) R1–R13.

[25] Jain A K, Duin R P W and Mao J , " Statistical pattern recognition: a review", *IEEE Trans. Pattern Anal. Mach.Intell.* 22,(2000) 4–37

[26] Lemm, S.; Schafer, C. and Curio, G., "BCI competition 2003-data set III: probabilistic modeling of sensorimotor mu- rhythms for classification of imaginary hand movements", *Biomedical Engineering, IEEE Transactions,* 51, (2004) 1077 – 1080.

[27] P. Lingras, C. Butz, " Rough set based 1-v-1 and 1-v-r approaches to support vector machine multi classification", *Inform. Sci.* 177 (2007) 3782–3798.

[28] BCI7. <www.dpmi.tu-graz.ac.at/schloegl/bci/bci7/>

Vickneswaran Jeyabalan obtained his BIT (Hons) degree in Software Engineering from Multimedia University, Malaysia and is pursuing his MEngSc also at the Multimedia University, Malaysia. His special fields of interest include EEG signal processing, Brain Machine Interfaces and Neural Networks. He has published papers related to Motor Imaginary based Brain Machine Interface in refereed Journals and International Conferences on Signal Processing. He has also received awards from IBM and participated in National level Microsoft Corp. sponsored competition for his research software applications.

Andrews Samraj is currently a lecturer at Multimedia University, Malaysia. His research interests are in the area of Bio-medical signal processing, Bio - signal processing, Bio-metrics, AI and Soft-computing, Rehabilitation Engineering and Parallel system architecture. Andrews obtained his MSc, MPhil, PGDPM from Bharathidasan Univierisity and Bharathiyar University, India. He has published many papers in referred journals and conferences.

Loo CK. Is currently the Dean of the Faculty of Information Science and Technology at Multimedia University, Malaysia, He received his Bachelor of Mechanical Engineering with First Class Honours from Universiti Malaya, Malaysia and PhD from University of Science, Malaysia. His research interests include Quantum Neural Networks, Synergetic Computing, Active force Control, Swarm Intelligence, Robotics and Intelligent Systems. He holds various research grants and has published many research papers in international conferences and journals.