We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS

BOOK CITATION INDEX

INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Using Autoregressive Models of Wavelet Bases in the Design of Mental Task-Based BCIs

Farshad Faradji, Farhad Faradji, Rabab K. Ward and
Gary E. Birch

Additional information is available at the end of the chapter

1. Introduction

1.1. Wavelet packet analysis

A powerful tool for analyzing the characteristics of the signal in the frequency domain as well as in the time domain is wavelet analysis. To analyze the signal, it can be decomposed into some levels successively by wavelet transform. At each level, decomposition yields two types of components: the approximations component which is the low-frequency high-scale portion of the signal, and the details component which is the high-frequency low-scale portion. The resultant approximations component is decomposed repetitively after each level. This is usually referred to as wavelet decomposition. Fig. 1.a shows a signal decomposed into three levels by wavelet decomposition.
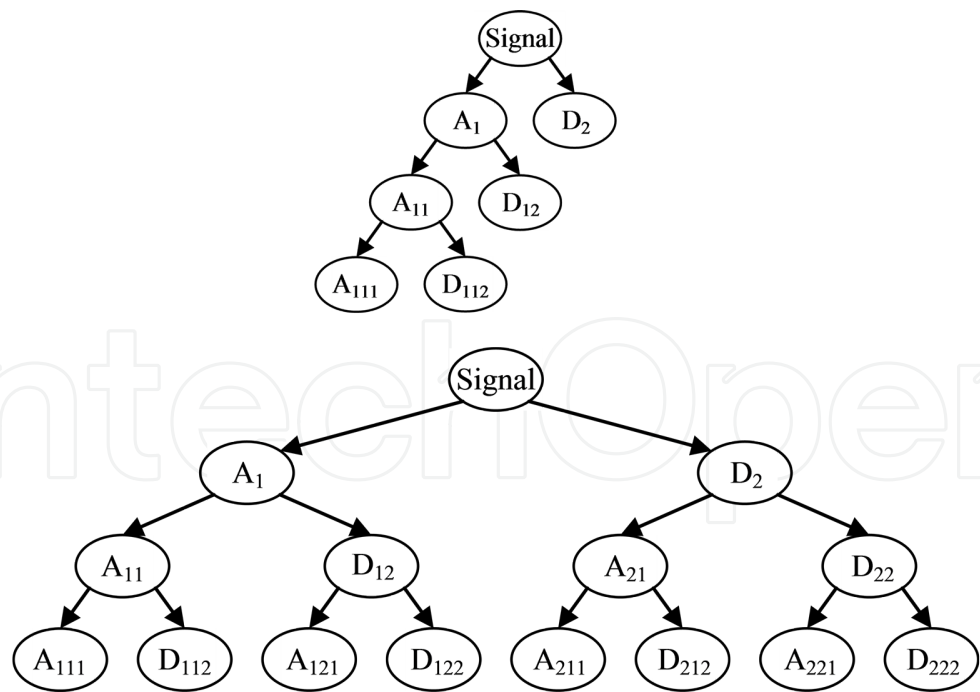
There is another approach for signal decomposition in which at each level, not only the approximations component but the details component is also decomposed. This is called wavelet packet analysis, and each of the components at different levels is referred to as a node or wavelet packet. Wavelet packet analysis is more flexible but more complicated than wavelet decomposition. Please see Fig. 1.b for more details. In this figure, the signal is decomposed into three levels using wavelet packet analysis.

Let us assume that the maximum frequency of the signal is f_m. At each level of decomposition, the approximations component has the lower half of the frequency spectrum of the signal decomposed, and the details component has the higher half of the signal's frequency spectrum. The frequency spectrums pertaining to different packets in three-level wavelet packet analysis is shown in Table 1.

Signal		Frequency $0-f_m$						
Level **1**	$0-0.5f_m$				$0.5f_m-f_m$			
2	$0-0.25f_m$		$0.25f_m-0.5f_m$		$0.5f_m-0.75f_m$		$0.75f_m-f_m$	
3	$0-0.125f_m$	$0.125f_m-0.25f_m$	$0.25f_m-0.375f_m$	$0.375f_m-0.5f_m$	$0.5f_m-0.625f_m$	$0.625f_m-0.75f_m$	$0.75f_m-0.875f_m$	$0.875f_m-f_m$

Table 1. Frequency Spectrums in Wavelet Packet Analysis

The packets produced by the wavelet packet analysis make an upside-down "tree", whose root is the initial signal and its branches are coming down. If we cut some branches of this tree (i.e., we exclude some packets from this tree), what remains is called a "sub-tree". The final nodes (terminal nodes or leaves) of each sub-tree are a "wavelet basis" for the initial signal. In other words, a wavelet basis is by definition a set of packets containing all non-overlapping frequency components of the initial signal. The packets of a basis can be selected from different levels. Each wavelet basis can serve as a representative for the initial signal in the wavelet domain. Here are some examples of the basis with respect to Fig. 1.b: $b_1 = \{A_1, D_2\}$, $b_2 = \{A_1, A_{21}, D_{22}\}$, $b_3 = \{A_{111}, D_{112}, D_{12}, D_2\}$, $b_4 = \{A_{11}, A_{121}, D_{122}, A_{211}, D_{212}, D_{22}\}$. It can be seen that wavelet decomposition is a special case of wavelet packet analysis (b_3).

Figure 1. Decomposition of a signal: (a) wavelet decomposition, (b) wavelet packet analysis

The best sub-tree (or basis) is a sub-tree (or basis) which minimizes a specific cost function. Cost functions are usually defined based on an entropy function (e.g., Shannon's entropy

function) [1]. One can also define his or her own cost function. To find the best sub-tree, the costs of all nodes in the main tree are measured. Then, in the direction from leaf to root and at each level, the cost of the node and its resultant child nodes are compared to each other. If the sum of the child nodes' costs is higher than or equal to the parent node's cost, the child nodes are cut off from the tree; otherwise, the child nodes remain in the tree, and the cost of the parent's node will be updated (replaced) with the sum of the child nodes' costs. This parent node is in turn one of the child nodes for the upper level. The same comparison (with updated costs) is done for the upper levels until the root level is reached. What remains at the end is the best sub-tree and its leaves are the best basis for the initial signal.

1.2. Brain–computer interfaces

Brain–Computer Interfaces (BCIs) providing an alternative channel for communication and intended to be used by motor-disabled individuals can be categorized into two main classes. Synchronized BCIs are the first and earliest class, in which the user is able to activate the system during pre-defined time periods only. The second class is referred to as asynchronous or self-paced BCIs. These systems are more useful, since the user can activate them whenever he or she wishes. At each time instant, a self-paced BCI is either in the intentional-control (IC) state or in the no-control (NC) state. IC is the active state, and NC is the inactive state (i.e., the state during which the output of the system is inactive).

The performance of a self-paced BCI is usually evaluated by two rates: the true positive rate (TPR) and the false positive rate (FPR). TPR is considered as the rate of correctly classifying intentional-control states, while FPR is defined as the rate of misclassifying no-control states.

A neurological phenomenon is a set of specific features or patterns in signals produced by the brain. They arise due to brain activities to which these phenomena are time-locked. Different types of neurological phenomena include the activity of neural cells (ANCs), P300, brain rhythms such as the Mu, Beta, and Gamma rhythms, movement-related potentials (MRPs), slow cortical potentials (SCPs), visual evoked potentials (VEPs), steady-state visual evoked potentials (SSVEPs), and mental tasks (MTs). For a review of the field of BCIs, please refer to [2]-[7].

The dataset of Keirn and Aunon [8] containing the EEG signals of five mental tasks is used in this paper, as it has been used in a variety of studies such as [9]-[55]. Although most of these studies can classify mental tasks to some extent, only a few of them care about false positives or confusion matrices and report them [32],[45]-[48]. The classification error is reported in studies [21],[31],[49] as well. The rates of correct classification are the only measure well paid attention to in the remaining studies. False positives are of great importance in BCI applications, they are hence fairly considered in this study.

1.3. Autoregressive modeling

The autoregressive (AR) model of an order K of a signal is defined as follows:

$$x[m] = \sum_{k=1}^{K} a_k x[m-k] + e[m] \qquad (1)$$

where $x[m]$ is the one-dimensional signal at time instant m, and a_k represents the AR coefficients. It is assumed that the error signal, $e[m]$, is a stochastic process and is independent of previous values of the signal x. It is also postulated that $e[m]$ has a zero mean and a finite variance. The autoregressive coefficients, a_k, should be estimated from the finite samples of the signal $x[m]$.

The most popular method to find the AR coefficients is the Burg algorithm [56] which computes coefficients at successive orders in the forward direction as well as in the backward direction. This method is used in this paper to estimate the AR coefficients.

Figuring out the optimal AR model order is not straightforward, even though some techniques including the Reflection Coefficient [18], the Information Theoretic criterion or Akaike Information Criterion (AIC), the Autoregressive Transfer Function criterion, and the Final Prediction Error (FPE) criterion [57] have been introduced. If the order of the model is too low, the whole signal cannot be captured in the model. On the other hand, the more the order is, the more portion of the noise is captured. Since there is no guarantee that the above mentioned techniques work well in every application and since FPR is of great importance in our application, we do not use these techniques to find the optimal AR model order. Instead, to be on the safe side, we vary the order in a reasonably large range and select the best order based on the performance of the system evaluated via nested five-fold cross-validation.

1.4. Quadratic discriminant analysis

The Quadratic Discriminant Analysis (QDA) classifier [58] assumes the classes have normal distributions. For QDA method of classification, unlike the linear discriminant analysis, the covariance matrices of the classes can be different. For a two-class problem, the quadratic discriminant function by definition is as follows:

$$qdf(x) = -\frac{1}{2}x^T(\hat{\Sigma}_1^{-1} - \hat{\Sigma}_2^{-1})x + (\hat{\mu}_1^T\hat{\Sigma}_1^{-1} - \hat{\mu}_2^T\hat{\Sigma}_2^{-1})x - \frac{1}{2}\ln\left(\frac{|\hat{\Sigma}_1|}{|\hat{\Sigma}_2|}\right) - \frac{1}{2}(\hat{\mu}_1^T\hat{\Sigma}_1^{-1}\hat{\mu}_1 - \hat{\mu}_2^T\hat{\Sigma}_2^{-1}\hat{\mu}_2) - \ln\left(\frac{C_{21}}{C_{12}}\frac{\pi_2}{\pi_1}\right) \qquad (2)$$

where x is the vector to be classified, $\hat{\mu}_1$, $\hat{\mu}_2$ are the estimated mean vectors of classes 1 and 2, $\hat{\Sigma}_1$, $\hat{\Sigma}_2$ are the estimated covariance matrices of class 1 and class 2, π_1, π_2 are the prior probabilities of the two classes, C_{12} is the cost of misclassifying a member of class 1 as class 2, and C_{21} is the cost due to misclassifying a member of class 2.

The decision rule for classification is:

$$x_0 \in \begin{cases} \omega_1 & \text{if } qdf(x_0) \geq 0 \\ \omega_2 & \text{if } qdf(x_0) < 0 \end{cases} \qquad (3)$$

where ω_1, ω_2 represent class 1 and class 2, respectively.

In this paper, the same value for the cost of false negative C_{21} and false positive C_{12} is used. It is also assumed that the a-priori probabilities of the two classes are equal.

2. Methods

2.1. Data

As mentioned in the Section 1.2, we used the EEG data which has been collected previously by Keirn and Aunon [8]. The EEG signals of this dataset belong to seven subjects, each performing five different mental tasks. The mental tasks include baseline, mentally computing a nontrivial multiplication, composing a letter to a friend mentally, rotating a three-dimensional object mentally, and visualizing writing a sequence of numbers on a blackboard. The subjects did not vocalize or gesture in any way when their signals were being recorded. Each session of recording comprises five trials of each mental task; therefore, there are a total of twenty five trials in a session. Only one session was performed on a single day. The length of each trial is ten seconds. Two of the subjects (Subjects 2 and 7) completed only one session, while one of them (Subject 5) completed three sessions. In this study, we used the data of subjects who completed at least 10 trials. Since the signals of Subject 4 were missing some data, we did not use them. New numbers were assigned to the subjects whose signals have been used in this study (see Table 2). Table 2 also shows the number of trials completed by each subject.
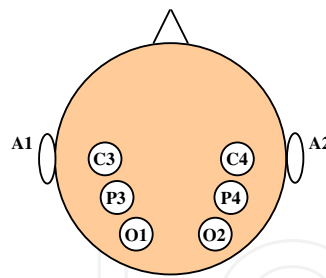
Subject number	original study	1	2	3	4	5	6	7
	this study	1	—	2	—	3	4	—
Number of completed trials		10	5	10	10	15	10	5

Table 2. The Number of Completed Trials for Each Subject

EEG signals were recorded from six channels (electrodes) while the subjects were seated in a room which is sound-controlled and with dim lighting. The electrodes were placed at positions C3, C4, P3, P4, O1, and O2 (based on the International 10-20 System) on the scalp. The reference electrodes were two electrically linked mastoids, A1 and A2. Fig. 2 shows the electrodes' locations. During recordings, the impedances between each electrode and the reference electrodes were kept below 5 kΩ. The signals were sampled at 250 Hz with an A/D converter (twelve-bit Lab Master) and a bank of amplifiers (Grass 7P511, with the band-pass filters set at 0.1-100 Hz). The system was calibrated at the beginning of each session with a known voltage.

Two EOG electrodes were placed below and at the outside corner of the left eye for detecting ocular artifacts. Since in this study, we did not remove any segment of the EEG signals due to ocular artifacts, the signals of EOG electrodes were not used.

Figure 2. Electrodes positions based on International 10-20 System.

Even though this dataset has not been collected in a self-paced paradigm, we are using it as an introductory exploration. It is obvious that in a self-paced paradigm, brain activities do not change, but since the pacing information (the exact start and end time of the mental tasks) is not known, training the BCI system would be more complicated.

2.2. Procedure

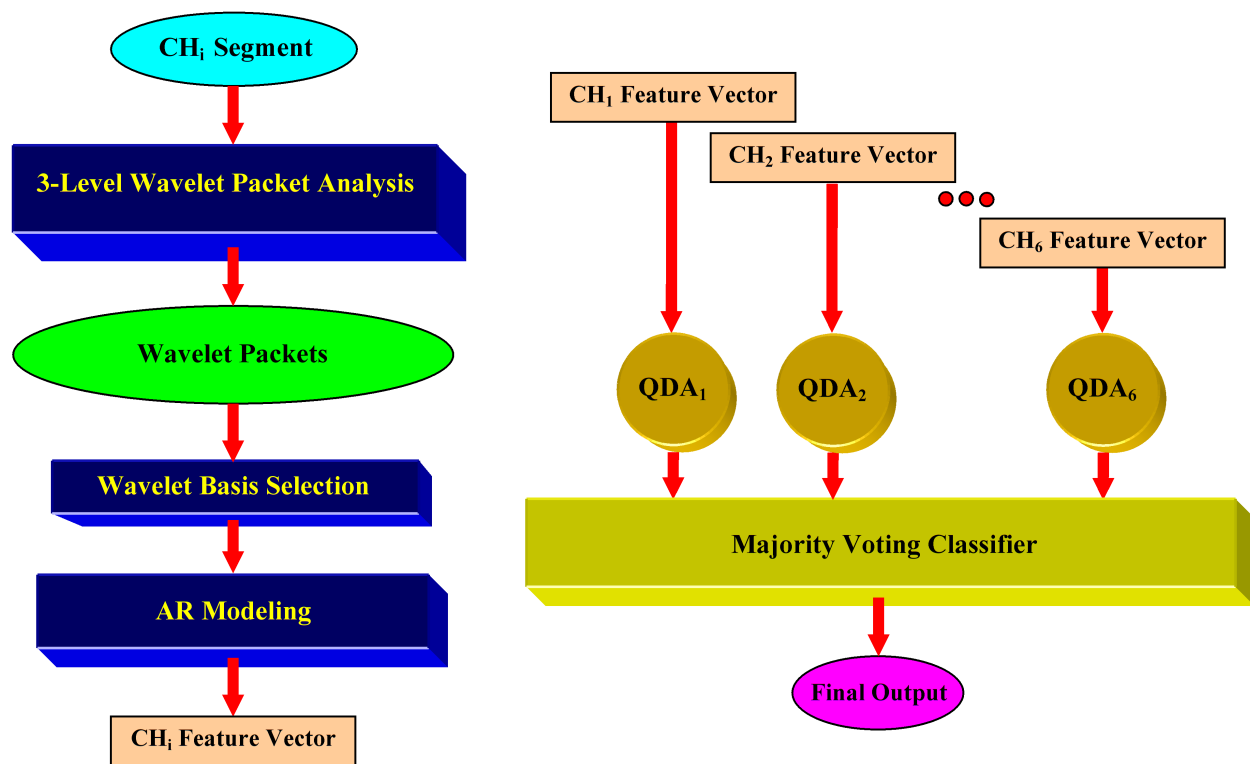### 2.2.1. The design of BCI systems

In this paper, we apply wavelet packet analysis to the design of a two-state self-paced mental task-based BCI. We develop and custom design five different BCIs for each subject based on the five mental tasks. In each BCI, one mental task is considered as the intentional-control task and the other four mental tasks are considered as the no-control tasks. Unlike the no-control tasks, the intentional-control task should activate the BCI system. Even though a BCI in which the baseline is the intentional-control task is practically useless, we consider it here for comparison purposes. We then determine the two most discriminatory mental tasks for each subject by comparing the performance of the five BCI systems of that subject. The overview of the proposed BCI system is illustrated in Fig. 3.

We customize the BCIs for every subject and mental task, since it has been proven that customized BCIs yield better results than general BCIs [59]-[60].

The EEG signals of four subjects are exploited. We use the first ten trials for Subject 3, and all ten trials for the other three subjects. The sampling rate is 250 Hz and each trial is 10 seconds long; therefore, there are 2500 samples in every trial of each mental task.

Each trial is divided into 45 256-sample overlapping segments. Each segment overlaps with the adjacent segment by 206 samples (about 80%). Hence, for each subject, the total number of segments for each mental task is 450. The segments with the length of more than 1 second are sufficiently long to get a good characteristic of the signal [61].

For each BCI, we have six different EEG channels. Each channel has its own feature vector and classifier. For each channel, the EEG segment is decomposed using wavelet packet analysis and the AR models of the resultant wavelet packets are estimated using the Burg algorithm. The AR coefficients of the packets belonging to a given wavelet basis are concatenated into a vector to form the channel's feature vector. The classifier of each channel is QDA. Each QDA

Figure 3. Overview of the proposed BCI system.

classifies the input EEG segment as an IC or NC segment. The task of the second-stage classifier which is a simple majority voting classifier is to determine whether the final output of the system is IC or NC.

2.2.2. Training, cross-validation, and testing

For each subject, the 5×450 segments pertaining to the five mental tasks are divided into a training set, a validation set, and a test set. The training set is used to train the system. The validation set is used to select the best wavelet, the best wavelet basis, and the best AR model order. The test set is used to evaluate the final performance of the system. The performance of a system evaluated based on a fixed split of data into training, validation and test sets is not accurate and robust, therefore, we perform nested five-fold (or 5×5) cross-validation. While the model selection is done during the inner cross-validation process, the system performance is estimated in the outer cross-validation.

The data are split into the five outer folds, for each of which 80% of the data is used for training and validation and 20% of the data is used for testing. The portion of data which is assigned for training and validation are further divided into five inner folds. For each inner fold, 80% of the data is used for training and the rest is used for validation. Hence, what we report as the cross-validation and testing results are the average over 25 and 5 different cases, respectively.

2.2.3. Different wavelet bases

Each of these segments is decomposed by wavelet packet analysis into three levels. In three-level wavelet packet decomposition, there exist 14 packets which can be seen in Fig. 1.b. We have 25 different wavelet bases representing the initial segment in the wavelet domain. These bases are listed in Table 3.

Basis	Packets							
1	A_1				D_2			
2	A_1				A_{21}	D_{22}		
3	A_1				A_{21}	A_{221}	D_{222}	
4	A_1				A_{211}	D_{212}	D_{22}	
5	A_1				A_{211}	D_{212}	A_{221}	D_{222}
6	A_{11}	D_{12}			D_2			
7	A_{11}	D_{12}			A_{21}	D_{22}		
8	A_{11}	D_{12}			A_{21}	A_{221}	D_{222}	
9	A_{11}	D_{12}			A_{211}	D_{212}	D_{22}	
10	A_{11}	D_{12}			A_{211}	D_{212}	A_{221}	D_{222}
11	A_{11}	A_{121}	D_{122}		D_2			
12	A_{11}	A_{121}	D_{122}		A_{21}	D_{22}		
13	A_{11}	A_{121}	D_{122}		A_{21}	A_{221}	D_{222}	
14	A_{11}	A_{121}	D_{122}		A_{211}	D_{212}	D_{22}	
15	A_{11}	A_{121}	D_{122}		A_{211}	D_{212}	A_{221}	D_{222}
16	A_{111}	D_{112}	D_{12}		D_2			
17	A_{111}	D_{112}	D_{12}		A_{21}	D_{22}		
18	A_{111}	D_{112}	D_{12}		A_{21}	A_{221}	D_{222}	
19	A_{111}	D_{112}	D_{12}		A_{211}	D_{212}	D_{22}	
20	A_{111}	D_{112}	D_{12}		A_{211}	D_{212}	A_{221}	D_{222}
21	A_{111}	D_{112}	A_{121}	D_{122}	D_2			
22	A_{111}	D_{112}	A_{121}	D_{122}	A_{21}	D_{22}		
23	A_{111}	D_{112}	A_{121}	D_{122}	A_{21}	A_{221}	D_{222}	
24	A_{111}	D_{112}	A_{121}	D_{122}	A_{211}	D_{212}	D_{22}	
25	A_{111}	D_{112}	A_{121}	D_{122}	A_{211}	D_{212}	A_{221}	D_{222}

Table 3. Different Wavelet Bases

2.2.4. The relationship between the packet length and the AR order

The length of a child packet is almost half of its parent packet's length. Since the initial decomposed segment is 256 samples long, the packets at levels 1, 2, and 3 contain approximately 128, 64, and 32 samples, respectively. When we estimate the packets' AR model, we

consider their lengths as a factor in selecting the appropriate AR order. In other words, we try to keep the same ratios between the orders of the packets at different levels. Therefore, if we set the AR order of a first-level packet as K, the AR orders for the packets at levels 2 and 3 would be close to $K/2$ and $K/4$, respectively. Table 4 provides the information on the sets of AR orders used for different levels of decomposition.

		AR Orders' Set												
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
	1	12	13	14	15	16	17	18	19	20	21	22	23	24
Level	**2**	6	7	7	8	8	9	9	10	10	11	11	12	12
	3	3	4	4	4	4	5	5	5	5	6	6	6	6

Table 4. The Sets of AR Orders Used for Different Levels of Decomposition

2.2.5. Different wavelets

Wavelets from various families are used. For each subject and each mental task, the wavelet with the best performance during nested five-fold cross-validation is selected. The 36 wavelets tested are from the Haar, Daubechies, Biorthogonal, Coiflets, and Symlets families. We assign a number to each of the wavelets and list them in Table 5.

Number	Wavelet	Family	Number	Wavelet	Family	Number	Wavelet	Family
1	db1 (Haar)		11	bior1.3		25	coif1	
2	db2		12	bior1.5		26	coif2	
3	db3		13	bior2.2		27	coif3	Coiflets
4	db4		14	bior2.4		28	coif4	
5	db5	Daubechies	15	bior2.6		29	coif5	
6	db6		16	bior2.8		30	sym2	
7	db7		17	bior3.1	Biorthogonal	31	sym3	
8	db8		18	bior3.3		32	sym4	
9	db9		19	bior3.5		33	sym5	Symlets
10	db10		20	bior3.7		34	sym6	
			21	bior3.9		35	sym7	
			22	bior4.4		36	sym8	
			23	bior5.5				
			24	bior6.8				

Table 5. Wavelets Used from Different Families

2.2.6. The proposed method to find the best wavelet basis

As mentioned earlier in the Introduction section, to find the best wavelet basis, different cost functions, which are usually different types of entropy functions, have been proposed. Unfortunately, the methodology based on the cost function is not working here because of two main reasons. First of all, since we have a number of segments for training the system which are not fully stationary, different wavelet bases are chosen for different segments. Hence, we can not come up with a single basis. Secondly, if we suppose that we are able to find the best basis based on a defined cost function, there is no guarantee that the selected basis has the best performance for our system. Therefore, we propose a method to find the best wavelet basis which does not have the above problems. The idea is very simple. We select a basis as the best basis which shows the best performance during a cross-validation process. There are two measures for performance evaluation of our system, TPR and FPR. The ratio of TPR to FPR is calculated to compare the performance with different wavelet basis. This method is not only applicable to the BCI system but also to any system which is based on classification.

2.2.7. Selecting the best wavelet and the best wavelet basis

For each subject and each mental task, we run a nested five-fold cross-validation process with the first set of AR model orders in Table 4 (i.e., 12, 6, and 3 for levels 1, 2, and 3, respectively) in order to find the best wavelet and the best wavelet basis. For each of the 36 wavelets, we test all 25 possible wavelet bases. The best wavelet basis is firstly determined for every wavelet based on the ratio of TPR/FPR. Secondly, we compare the performance of the system with the best bases of different wavelets, and figure out the wavelet whose best basis yields the best system performance (in terms of the ratio of TPR to FPR). It is noteworthy that the results would be the same if we first find the best wavelet for each of the 25 wavelet bases and then select the basis with the best performance.

2.2.8. Optimizing the AR model order

Having selected the best wavelet and wavelet basis for the BCI of each subject and each mental task, we find the optimal AR order via another nested five-fold cross-validation process. To this end, we test different sets of AR orders (i.e., sets 2, 3, …, 13 in Table 4) using the best wavelet and the best basis as previously determined, and select the set with the highest TPR/FPR ratio.

2.2.9. Testing the system

We test the system via the outer fold of nested five-fold cross-validation with the selected wavelet, the best wavelet basis, and the optimal AR order set for every BCI belonging to each subject and mental task. The results are given in the next section.

3. Results

The results of the cross-validation process (at AR order set 1 and the optimal AR order set) and the results of testing (at the optimal AR order set) are summarized in Table 6. This table

also shows the selected wavelet and the best wavelet basis for each BCI. The performance of a system which is based on three-level wavelet decomposition (instead of wavelet packet analysis) is furthermore given in Table 6 for comparison purposes.

		Baseline					Multiplication					Letter Composing					Rotation					Counting				
		AR Order Set	TPR Mean	TPR SD	FPR Mean	FPR SD	AR Order Set	TPR Mean	TPR SD	FPR Mean	FPR SD	AR Order Set	TPR Mean	TPR SD	FPR Mean	FPR SD	AR Order Set	TPR Mean	TPR SD	FPR Mean	FPR SD	AR Order Set	TPR Mean	TPR SD	FPR Mean	FPR SD
Subject 1	Validation	1	70.89	3.62	2.92	0.30	**1**	**71.17**	**1.06**	**3.18**	**0.37**	1	66.33	1.61	2.35	0.33	1	70.00	1.71	3.44	0.38	1	69.39	0.54	3.40	0.08
	Validation	13	65.61	2.16	0.08	0.14	**13**	**53.67**	**1.87**	**0.19**	**0.14**	13	57.78	0.92	0.28	0.07	13	61.44	0.69	0.21	0.08	13	63.06	1.40	0.21	0.16
	Testing	13	65.56	8.99	0.17	0.25	**13**	**52.67**	**6.55**	**0.28**	**0.20**	13	61.33	3.37	0.33	0.12	13	62.44	3.46	0.17	0.15	13	64.89	5.70	0.39	0.46
	Wavelet	13					**3**					25					18					3				
	Best Basis	1					**1**					1					1					1				
	Basis 16	13	47.56	3.96	0.22	0.36	**13**	**42.22**	**5.72**	**0.72**	**0.32**	13	44.89	5.07	0.61	0.36	13	41.56	5.42	0.61	0.36	13	49.11	5.74	0.44	0.25
Subject 2	Validation	1	59.50	1.14	4.25	0.29	1	60.39	2.32	4.90	0.19	1	79.72	1.11	7.90	0.26	1	64.72	2.99	6.82	0.37	**1**	**68.39**	**2.97**	**6.75**	**0.61**
	Validation	13	56.22	3.62	0.38	0.13	13	56.94	2.05	0.24	0.03	13	68.28	2.66	0.60	0.11	13	45.28	2.69	0.42	0.16	**13**	**63.39**	**1.57**	**0.50**	**0.12**
	Testing	13	60.89	4.26	0.44	0.50	13	56.44	5.85	0.17	0.25	13	68.22	5.75	0.39	0.32	13	46.67	5.09	0.67	0.75	**13**	**64.00**	**5.70**	**0.22**	**0.12**
	Wavelet	32					2					2					9					**2**				
	Best Basis	1					1					1					1					**1**				
	Basis 16	13	46.89	2.88	0.89	0.63	13	44.67	4.33	0.78	0.36	13	52.89	5.01	1.17	0.80	13	44.22	4.04	0.67	0.58	**13**	**46.89**	**4.40**	**1.11**	**0.56**
Subject 3	Validation	1	62.67	2.07	3.67	0.20	1	66.22	1.96	3.53	0.44	**1**	**66.56**	**2.82**	**2.51**	**0.15**	1	70.72	1.76	3.07	0.29	1	65.78	1.77	3.96	0.41
	Validation	13	60.72	0.65	0.19	0.12	13	59.89	1.48	0.11	0.08	**13**	**60.50**	**1.09**	**0.13**	**0.11**	13	52.39	2.31	0.14	0.06	13	61.94	1.62	0.03	0.09
	Testing	13	58.00	5.52	0.17	0.25	13	58.67	5.18	0.17	0.15	**13**	**60.00**	**3.24**	**0.00**	**0.00**	13	51.33	7.71	0.06	0.12	13	59.56	6.79	0.00	0.00
	Wavelet	11					34					**2**					5					2				
	Best Basis	1					1					**1**					1					1				
	Basis 16	13	37.78	6.89	0.72	0.37	13	45.33	2.65	0.22	0.12	**13**	**43.78**	**4.27**	**0.17**	**0.15**	13	48.22	5.86	0.11	0.15	13	39.11	4.26	0.50	0.30
Subject 4	Validation	1	73.67	4.36	3.89	0.22	1	79.00	1.61	3.35	0.20	**1**	**70.78**	**1.27**	**4.76**	**0.52**	1	86.89	1.64	4.14	0.47	1	70.83	1.08	3.04	0.26
	Validation	13	67.17	2.58	0.26	0.12	13	70.83	2.29	0.13	0.12	**13**	**62.06**	**3.11**	**0.10**	**0.20**	13	73.94	0.63	0.28	0.16	13	61.61	1.07	0.13	0.06
	Testing	13	68.89	3.24	0.28	0.39	13	69.78	2.41	0.17	0.15	**13**	**64.44**	**6.71**	**0.11**	**0.25**	13	75.78	3.08	0.33	0.30	13	63.56	9.04	0.06	0.12
	Wavelet	4					25					**19**					14					5				
	Best Basis	1					1					**1**					1					1				
	Basis 16	13	57.11	6.31	0.72	0.50	13	57.33	2.30	0.39	0.32	**13**	**48.00**	**3.08**	**0.50**	**0.46**	13	65.33	6.64	0.61	0.23	13	47.11	6.92	0.39	0.25

Table 6. Performance of BCIs for Different Subjects and Tasks during Cross-Validation and Testing

Table 6 is divided into 20 sections. Each section contains the information about the BCI belonging to a specific subject and mental task. The first line in every section shows the result of the first cross-validation process at AR order set 1. As mentioned before, this cross-validation is performed in order to select the best wavelet and the best wavelet basis for each BCI. The results include the mean and the standard deviation of the TPR and FPR values for the selected case (i.e., the best wavelet and basis). The best wavelet and the best basis are given on lines four and five of the table, respectively.

The optimal AR model order set is determined based on the results of the second cross-validation process done with the selected wavelet and basis. The results of the second cross-validation with the optimal AR order make the second lines of sections. It is noteworthy that the optimal AR order set for all BCIs is the last set which has the largest numbers.

The performance of the BCI systems with the best configuration (i.e., using the best wavelet, best basis, and optimal AR order set) are also given on the third line of each section. The last line of each section presents the performance of the BCI based on three-level wavelet decomposition as an example to be compared with the performance of the system based on the best

basis. As previously discussed, three-level wavelet decomposition is one of the existing bases in three-level wavelet packet analysis (basis 16 according to Table 3).

The most discriminatory task for each subject is determined by comparing the performance of five BCIs pertaining to the subject. The most discriminatory mental task is in a section with solid black borders. For each subject, the second most discriminatory task is also chosen based on the results and shown in bold in the table.

Table 7 presents the average system performance of the BCIs for each subject (over tasks) and for each task (over subjects).

The preliminary results have been published in [62].

			TPR		FPR	
			Mean	**SD**	**Mean**	**SD**
Average over Tasks	**Subject**	**1**	61.38	5.61	0.27	0.24
		2	59.24	5.33	0.38	0.39
		3	57.51	5.69	0.08	0.10
		4	68.49	4.90	0.19	0.24
Average over Subjects	**Task**	**Baseline**	63.34	5.50	0.27	0.35
		Multiplication	59.39	5.00	0.20	0.19
		Letter Composing	63.50	4.77	0.21	0.17
		Rotation	59.06	4.84	0.31	0.33
		Counting	63.00	6.81	0.17	0.18
Total Average			**61.66**	**5.38**	**0.23**	**0.24**

Table 7. Average Performance of BCI Systems

4. Discussion

4.1. Best wavelet bases

As seen from Table 6, the best wavelet basis for all subjects and mental tasks is surprisingly the first basis which is made of the approximations and details components of the first level of wavelet decomposition. This implies that for the proposed BCIs, it is enough to decompose the signal into the first level, and further decomposition degrades the system performance.

4.2. Most discriminatory tasks

To determine the most discriminatory tasks for each subject, we consider the FPR values during testing and cross-validation at the optimal AR order set. We put the main weight on the FPR

of testing since not only the portion of the dataset used for testing is larger than the portion used for cross-validation, but also the testing portion is completely separate from the portions exploited during training and cross-validation processes. The TPR values during testing and cross-validation are then considered if necessary to find out the most discriminatory tasks.

For Subject 1, the testing FPR for the baseline and the rotation tasks are the same and the lowest. The cross-validation FPR is lower for the baseline. Hence, the most discriminatory task must be the baseline. Since a BCI based on the baseline is activated when the subject wishes to relax and think of nothing, it is practically useless; therefore, we do not consider the baseline as the most discriminatory task. The rotation task is then selected as the most discriminatory task. The second most discriminatory task for Subject 1 is the multiplication.

For Subject 2, the most discriminatory task is the multiplication since it has the lowest testing and cross-validation FPRs. The counting is the second most discriminatory task for this subject.

For Subjects 3 and 4, the most and the second most discriminatory tasks are the counting and the letter composing, respectively. For Subject 3, the FPR values for these two tasks interestingly reach zero during testing.

4.3. Selected wavelets

Unlike the basis, the selected wavelets are not the same for all subjects and mental tasks. In eleven BCIs (out of the twenty BCIs designed for different subjects and tasks), a wavelet from the Daubechies family has been chosen as the best wavelet. Wavelet 'db2', the most selected wavelet, is the best wavelet for five BCIs. The BCIs for the most and the second most discriminatory tasks of Subjects 2 and 3 has the best performance with this wavelet among all other wavelets. In five BCIs, the Biorthogonal family is the best. Each of the Coiflets and Symlets families are also selected in two BCIs.
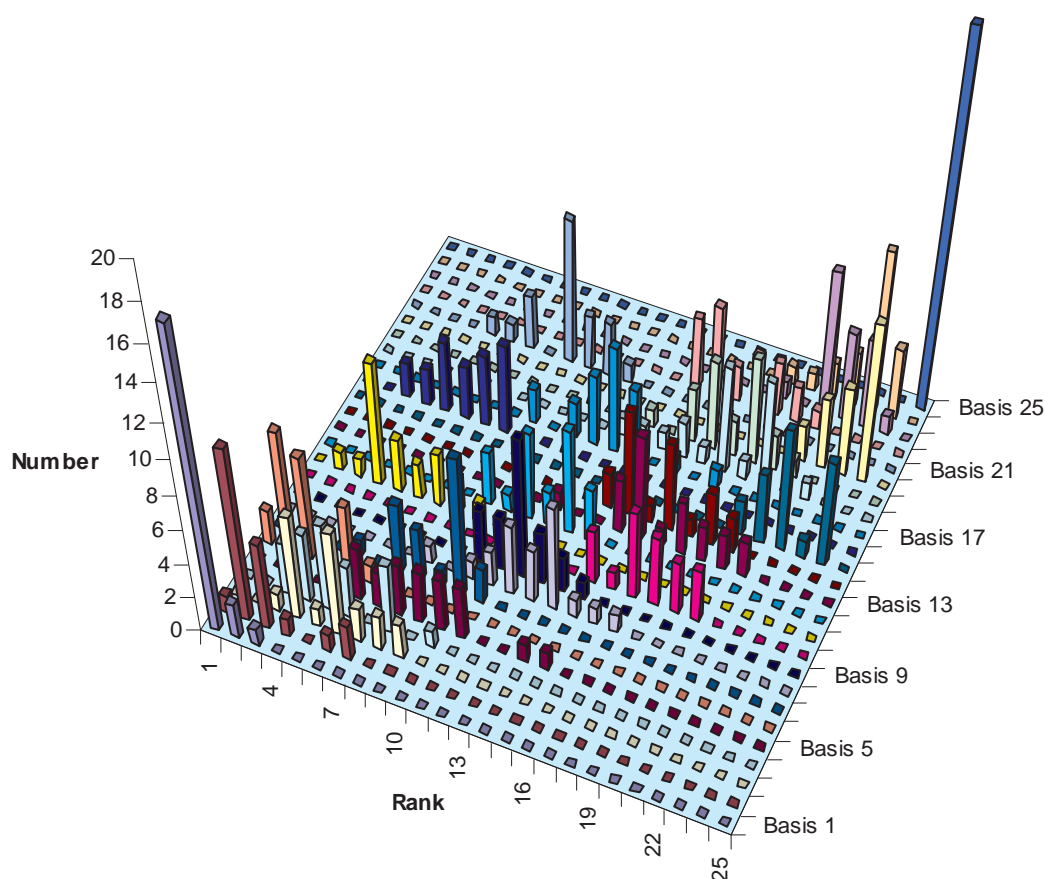
4.4. Average system performance

According to Table 7, the BCIs of Subject 3 have generally the best performance, with the average FPR of 0.08% and the average TPR of 57.51%. Moreover, the BCIs based on the counting task have the best performance overall. The average performance of BCIs based on the counting has FPR and TPR values of 0.17% and 63.00%.

4.5. Comparison of different wavelet bases

For each BCI, we consider and evaluate the system performance with different wavelet bases as further analysis. We then sort and rank the bases based on the ratio of TPR to FPR during testing. Considering the performance of each wavelet basis for different subjects and mental tasks, we count the number of cases that each basis ranks n-th among the other bases. The results are summarized in Table 8. The corresponding three-dimensional histograms are shown in Fig. 4. In this figure, the horizontal axes are related to different wavelet bases and their ranks. The vertical axis is showing the number of times that a basis has a specific rank amongst other bases. It can be seen from this bar diagram that as the basis number is increasing,

the ranks of the basis is getting worse, i.e., the first basis has the highest ranks and the last basis has the worst rank. The results are almost along with the cross-validation results (with the first set of AR orders) and show that the first basis is the best basis for all BCIs except for three of them belonging to the multiplication task of Subject 1 and the rotation task of Subjects 2 and 3. For the multiplication task of Subject 1, the best basis is basis 2. The basis 6 ranks first for the rotation task of Subjects 2 and 3. The rank of the first basis for the multiplication task of Subject 1 and the rotation task of Subject 2 is two. The first basis is ranked third for the rotation task of Subject 3. In all other 17 BCIs, the first basis is the best.

Figure 4. Comparison of different wavelet bases.

Number of Occurrence / **Rank**

	Wavelet Basis																								
Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
1	17	1				2																			
2	2	10				7					1														
3	1	5	1	3		6	1				1					2									
4		1	6	4							7					2									
5			1	5	1	4	1				3					4					1				
6		1	6	3	3	1					2					3					1				
7		2	2		2		4				3					4					3				
8			2	4	3		3	1			2					5									
9		2		3		1					1	3				2					8				
10			1	3				8	1	3	1										3				
11				3				2	2	3	5					2					3				
12								4	8		2	1				4					1				
13								3	3	1		6					6	1							
14						1		6	2			3		2			4	1	1						
15						1		1	1	3			3	6				1				4			
16								1		1			6	1			1	2	3			5			
17								1		5			1	5				1	5			2			
18										4			3	1				1	6	2		3			
19										3			2	3					1	6		3	1	1	
20										3			2	2	2				6	2		2		1	
21													2		4					1	2	1	8	2	
22															7			1		4			5	3	
23															1					5			5	9	
24															6					9			1	4	
25																									20

Table 8. Comparing Different Wavelet Bases during Testing at AR Order Set 13

5. Conclusion

In this paper, we presented a method to select the best wavelet basis in the design of a two-state self-paced mental task-based BCI. The use of the proposed methodology is not limited to the BCI systems and it can be also used in other applications. The previously introduced methods (based on cost functions) to find the best wavelet basis generally has two major drawbacks. First of all, they are not practical in classification problems where we have different training signal segments and each of them may result in a different basis. In this case, we cannot

come up with a unique basis for all of the training signal segments; therefore, we are not able to finalize our design of the classification system. The second drawback is resulted by this fact that, supposing we can find a unique best wavelet basis for all training segments, there is no guarantee that this wavelet basis yields the best classification accuracy. Because of these two reasons, we decided to propose a method based on the classification accuracy to find the best basis. Since our BCI systems are evaluated by true positive and false positive rates, we used these two measures in the process of finding the best wavelet basis. It is worth noting that any other kind of classification accuracy measure can be potentially exploited in our proposed method to select the best basis.

We have tested our proposed method in the design of mental task-based BCIs. The output of the BCI should be activated when the subject performs a specific mental task. The aim is to minimize and maximize the false activation rate and the true activation rate, respectively.

The scalar autoregressive model coefficients of the components of the best wavelet basis were used as features. The classifiers studied were based on QDA and majority voting. We performed nested five-fold cross-validation two times to choose the best wavelet, the best wavelet basis, and the best autoregressive model orders. Results have shown that the most discriminatory tasks are different amongst the subjects confirming the findings of previous studies ([20], [30], [33], and [42]) based on the same dataset.

For each subject and each mental task, the best configuration (i.e., the best wavelet, the best wavelet basis, and the optimum AR order) is found during offline analysis of the data. During online analysis, the best configuration is used. Therefore, the system is applicable to real-time applications.

Author details

Farshad Faradji[1], Farhad Faradji[2,3*], Rabab K. Ward[3] and Gary E. Birch[3]

1 Qom University, Qom, Iran

2 Amirkabir University of Technology, Tehran, Iran

3 University of British Columbia, Vancouver, British Columbia, Canada

References

[1]   R.R. Coifman and M.V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. on Inf. Theory*, vol. 38, no. 2, pp. 713–718, Mar. 1992.

[2]   M.A. Nicolelis, "Brain-machine interfaces to restore motor function and probe neural circuits," *Nat. Rev. Neurosci.*, vol. 4, no. 5, pp. 417–422, 2003.

[3] S.G. Mason and G.E. Birch, "A general framework for brain-computer interface design," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, pp. 70–85, Mar. 2003.

[4] T.M. Vaughan, "Guest editorial brain-computer interface technology: A review of the second international meeting," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 94–109, Jun. 2003.

[5] J.R. Wolpaw, "Brain-computer interfaces (BCIs) for communication and control: a mini-review," *Supplements to Clin. Neurophysiol.*, vol. 57, pp. 607–613, 2004.

[6] S.G. Mason, A. Bashashati, M. Fatourechi, K.F. Navarro, and G.E. Birch, "A comprehensive survey of brain interface technology designs," *Ann. Biomed. Eng.*, vol. 35, no. 2, pp. 137–169, Feb. 2007.

[7] A. Bashashati, M. Fatourechi, R.K. Ward, and G.E. Birch, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals," *J. Neural Eng.*, vol. 4, no. 2, pp. R35–57, Jun. 2007.

[8] Z.A. Keirn and J.I. Aunon, "A new mode of communication between man and his surroundings," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 12, pp. 1209–1214, Dec. 1990.

[9] Z.A. Keirn and J.I. Aunon, "Man-machine communications through brain-wave processing," *IEEE Eng. Med. Biol.*, pp. 55–57, Mar. 1990.

[10] C.W. Anderson, E. Stolz, and S. Shamsunder, "Discriminating mental tasks using EEG represented by AR models," *In Proc. 17th IEEE EMBS*, pp. 875–876, 1995.

[11] C.W. Anderson, S.V. Devulapalli, and E.A. Stolz, "Determining mental state from EEG signals using neural networks," *Sci. Program.*, vol. 4, no. 3, pp. 171–183, 1995.

[12] C.W. Anderson and Z. Sijerčić, "Classification of EEG signals from four subjects during five mental tasks," *In Proc. Conf. Eng. App. Neural Net.*, pp. 407–414, Jun. 1996.

[13] C.W. Anderson, "Effects of variations in neural network topology and output averaging on the discrimination of mental tasks from spontaneous electroencephalogram," *J. Intell. Syst.*, vol. 7, pp. 165–190, 1997.

[14] C.W. Anderson, E.A. Stolz, and S. Shumsunder, "Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks," *IEEE Trans. Biomed. Eng.*, vol. 45, no.3, pp. 277–286, Mar. 1998.

[15] R. Palaniappan, P. Raveendran, S. Nishida, and N. Saiwaki, "Evolutionary Fuzzy ARTMAP for autoregressive model order selection and classification of EEG signals," *In Proc. IEEE Int. Conf. Syst. Man Cybern.*, pp. 3682–3686, Oct. 2000.

[16] R. Palaniappan, P. Raveendran, S. Nishida, and N. Saiwaki, "Fuzzy Artmap classification of mental tasks using segmented and overlapped EEG signals," *In Proc. IEEE Region 10 Conf.*, pp. II-388–II-391, Sep. 2000.

[17] R. Palaniappan and P. Raveendran, "A new mode of EEG based communication," *In Proc. IEEE Int. Joint Conf. Neural Net.*, vol. 4, pp. 2679–2682, Jul. 2001.

[18] R. Palaniappan, P. Raveendran, S. Nishida, and N. Saiwaki, "Autoregressive spectral analysis and model order selection criteria for EEG signals," *In Proc. IEEE Region 10 Conf.*, pp. II-126–II-129, Sep. 2000.

[19] M.I. Bhatti, A. Pervaiz, and M.H. Baig, "EEG signal decomposition and improved spectral analysis using wavelet transform," *In Proc. 23rd IEEE EMBS*, pp. 1862–1864, Oct. 2001.

[20] R. Palaniappan, R. Paramesran, S. Nishida, and N. Saiwaki, "A new brain-computer interface design using fuzzy ARTMAP," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 10, no. 3, pp. 140–148, Sep. 2002.

[21] V.A. Maiorescu, M. Serban, and A.M. Lazar, "Classification of EEG signals represented by AR models for cognitive tasks - a neural network based method," *In Proc. Int. Symp. Signals Circuits Syst.*, vol. 2, pp. 441–444, 2003.

[22] D. Garrett, D.A. Peterson, C.W. Anderson, and M.H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 141–144, Jun. 2003.

[23] D. Liu, Z. Jiang, W. Cong, and H. Feng, "Detect determinism of spontaneous EEG with a multi-channel reconstruction method," *In Proc. IEEE Int. Conf. Neural Net. Signal Process.*, pp. 708–711, Dec. 2003.

[24] X. Wu and X. Guo, "Mental EEG Analysis based on Independent Component Analysis," *In Proc. 3rd Int. Symp. Image Signal Process. Anal.*, pp. 327–331, Sep. 2003.

[25] D. Liu, Z. Jian, and H. Feng, "Separating the different components of spontaneous EEG by optimized ICA," *In Proc. IEEE Int. Conf. Neural Net. Signal Process.*, pp. 1334–1337, Dec. 2003.

[26] J.Z. Xue, H. Zhang, C.X. Zheng, and X.G. Yan, "Wavelet packet transform for feature extraction of EEG during mental tasks," *In Proc. 2nd Int. Conf. Machine Learn. Cybern.*, pp. 360–363, Nov. 2003.

[27] G.A. Barreto, R.A. Frota, and F.N.S. de Medeiros, "On the classification of mental tasks: A performance comparison of neural and statistical approaches," *In Proc. IEEE Workshop Machine Learn. Signal Process.*, pp. 529–538, 2004.

[28] M.S. Daud and J. Yunus, "Classification of mental tasks using de-noised EEG signals," *In Proc. 7th Int. Conf. Signal Process.*, pp. 2206–2209, 2004.

[29] K. Tavakolian, S. Rezaei, and S.K. Setarehdan, "Choosing optimal mental tasks for classification in brain computer interfaces," *In Proc. Int. Conf. Artificial Intell. and App.*, pp. 396-399, Feb. 2004.

[30] K. Tavakolian and S. Rezaei, "Classification of mental tasks using Gaussian mixture Bayesian network classifiers," *In Proc. IEEE Int. Workshop Biomed. Circuits Syst.*, pp. S3.6-9–S3.6-11, Dec. 2004.

[31] R. Rao and R. Derakhshani, "A comparison of EEG preprocessing methods using time delay neural networks," *In Proc. 2nd Int. IEEE EMBS Conf. Neural Eng.*, pp. 262–264, Mar. 2005.

[32] R. Palaniappan, "Identifying individuality using mental task based brain computer interface," *In Proc. 3rd Int. Conf. Intell. Sensing Infor. Process.*, pp. 239–242, Dec. 2005.

[33] R. Palaniappan, "Brain computer interface design using band powers extracted during mental tasks," *In Proc. 2nd Int. IEEE EMBS Conf. Neural Eng.*, pp. 321–324, Mar. 2005.

[34] N. Huan and R. Palaniappan, "Classification of mental tasks using fixed and adaptive autoregressive models of EEG signals," *In Proc. 2nd IEEE EMBS Conf. Neural Eng.*, pp. 633–636, Mar. 2005.

[35] R. Palaniappan and N. Huan, "Improving the performance of two-state mental task brain-computer interface design using linear discriminant classifier," *In Proc. EUROCON*, vol. 1, pp. 409–412, Nov. 2005.

[36] S. Rezaei, K. Tavakolian, and K. Naziripour, "Comparison of five different classifiers for classification of mental tasks," *In Proc. 27th IEEE EMBS*, pp. 6007–6010, Sep. 2005.

[37] Z. Jiang, Y. Ning, B. An, A. Li, and H. Feng, "Detecting mental EEG properties using detrended fluctuation analysis," *In Proc. 27th IEEE EMBS*, pp. 2017–2020, Sep. 2005.

[38] M.-C. Setban and D.-M. Dobrea, "Discrimination between cognitive tasks - a comparative study," *In Proc. Int. Symp.* Signals Circuits Syst., pp. 805–808, Jul. 2005.

[39] H. Liu, J. Wang, and C. Zheng, "Mental tasks classification and their EEG structures analysis by using the growing hierarchical self-organizing map," *In Proc. 1st Int. Conf. Neural Interface Control*, pp. 115–118, May 2005.

[40] C. Gope, N. Kehtarnavaz, and D. Nair, "Neural network classification of EEG signals using time-frequency representation," In Proc. IEEE Int. Joint Conf. Neural Net., vol. 4, pp. 2502–2507, Aug. 2005.

[41] H. Liu, J. Wang, C. Zheng, and P. He, "Study on the effect of different frequency bands of EEG signals on mental tasks classification," *In Proc. 27th IEEE EMBS*, pp. 5369–5372, Sep. 2005.

[42] R. Palaniappan, "Utilizing gamma band to improve mental task based brain-computer interface design," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 3, pp. 299–303, Sep. 2006.

[43] G. Yan, X. Guo, R. Yan, and B. Yang, "Nonlinear quadratic phase coupling on EEG based on $1^{1/2}$-dimension spectrum," *In Proc. 3rd Int. Conf.* Advances Med. Signal Infor. Process., pp. 1–4, Jul. 2006.

[44] F. Abdollahi and A. Motie-Nasrabadi, "Combination of frequency bands in EEG for feature reduction in mental task classification," *In Proc. 28th IEEE EMBS*, pp. 1146–1149, Sep. 2006.

[45] K. Nakayama and K. Inagaki, "A brain computer interface based on neural network with efficient pre-processing," *In Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, pp. 673–676, Dec. 2006.

[46] C.W. Anderson, J.N. Knight, T. O'Connor, M.J. Kirby, and A. Sokolov, "Geometric subspace methods and time-delay embedding for EEG artifact removal and classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 142–146, 2006.

[47] D.-M. Dobrea and M.-C. Dobrea, "An EEG (bio) technological system for assisting the disabled people," *In Proc. 5th IEEE Int. Conf. Comput. Cybern.*, pp. 191–196, Oct. 2007.

[48] D.-M. Dobrea, M.-C. Dobrea, and M. Costin, "An EEG coherence based method used for mental tasks classification," *In Proc. 5th IEEE Int. Conf. Comput. Cybern.*, pp. 185–190, Oct. 2007.

[49] K. Nakayama Y. Kaneda, and A. Hirano, "A brain computer interface based on FFT and multilayer neural network - feature extraction and generalization," *In Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, pp. 826–829, Nov. 2007.

[50] L. Zhiwei and S. Minfen, "Classification of mental task EEG signals using wavelet packet entropy and SVM," *In Proc. 8th Int. Conf. Elec. Measure. Instr.*, pp. 3-906–3-909, Aug. 2007.

[51] B.T. Skinner, H.T. Nguyen, and D.K. Liu, "Classification of EEG signals using a genetic-based machine learning classifier," *In Proc. 29th IEEE EMBS*, pp. 3120–3123, Aug. 2007.

[52] F. Abdollahi, S.K. Setarehdan, and A.M. Nasrabadi, "Locating information maximization time in EEG signals recorded during mental tasks," *In Proc. 5th Int. Symp. Image Signal Process. Anal.*, pp. 238–241, Sep. 2007.

[53] C.R. Hema, M.P. Paulraj, R. Nagarajan, S. Yaacob, and A.H. Adom, "Fuzzy based classification of EEG mental tasks for a brain machine interface," *In Proc. 3rd Int. Conf. Intell. Infor. Hiding Multimed. Signal Process.*, vol. 1, pp. 53–56, Nov. 2007.

[54] S.M. Hosni, M.E. Gadallah, S.F. Bahgat, and M.S. AbdelWahab, "Classification of EEG signals using different feature extraction techniques for mental-task BCI," *In Proc. Int. Conf.* Computer Eng. Syst., pp. 220–226, Nov. 2007.

[55] M.P. Paulraj, C.R. Hema, R. Nagarajan, S. Yaacob, and A.H. Adom, "EEG classification using radial basis PSO neural network for brain machine interfaces," *In Proc. 5th Student Conf. Research Develop.*, pp. 1–5, Dec. 2007.

[56] J.P. Burg, "A new analysis technique for time series data," NATO Adv. Study Inst. on Signal Processing with Emphasis on Underwater Acoustics, Enschede, The Netherlands, Aug. 1968, reprinted in *Modern Spectrum Analysis*, D.G. Childers, ed., IEEE Press, pp. 42–48, New York, 1978.

[57] P.J. Franaszczuk, K.J. Blinowska, and M. Kowalczyk, "The application of parametric multichannel spectral estimates in the study of electrical brain activity," *Biological Cybern.*, vol. 51, pp. 239–247, 1985.

[58] A.C. Atkinson, M. Riani, and A. Cerioli, *Exploring Multivariate Data with the Forward Search*, Springer Series in Statistics, XXI, 621 p., 2004, ch. 6.

[59] G. Blanchard and B. Blankertz, "BCI competition 2003-dataset IIa: Spatial patterns of self-controlled brain rhythm modulations," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1062–1066, Jun. 2004.

[60] A. Bashashati, M. Fatourechi, R.K. Ward, and G.E. Birch, "User customization of the feature generator of an asynchronous brain interface," *Ann. Biomed. Eng.*, vol. 34, no. 6, pp. 1051–1060, Jun. 2006.

[61] G.E. Birch, P.D. Lawrence, J.C. Lind, and R.D. Hare, "Application of prewhitening to AR spectral estimation of EEG," *IEEE Trans. Biomed. Eng.*, vol. 35, no. 8, pp. 640–645, Aug. 1988.

[62] F. Faradji, R.K. Ward, and G.E. Birch, "A simple approach to find the best wavelet basis in classification problems," *In Proc. 20th Int. Conf. Pattern Recog.*, pp. 641–644, Aug. 2010.