A feature extraction technique of EEG based on EMD-BP for motor imagery classification in BCI

Dalila Trad
ESSTT, LaTICE
5, Avenue Taha Hussein,
B. P. : 56, Bab Menara, 1008 Tunisia
Email: trad.dalila@gmail.com,

Tarik Al-ani
UVSQ, LISV,
10/12 Avenue de l'Europe,
78140 Vélizy, France
and
ESIEE-Paris, Dep. Informatique,
Cité Descartes-BP 99,
93162 Noisy-Le-Grand, France
Email: t.alani@esiee.fr

Mohamed Jemni
ESSTT, LaTICE
5, Avenue Taha Hussein,
B. P. : 56, Bab Menara, 1008 Tunisia
Email: mohamed.jemni@fst.rnu.tn

***Abstract*—The aim of this paper is to investigate a nonlinear approach for feature extraction of *Electroencephalogram* (*EEG*) signals in order to classify motor imagery for *Brain Computer Interface* (*BCI*). This approach consists of combining the *Empirical Mode Decomposition* (*EMD*) and *band power* (*BP*). Considering the non-stationary and nonlinear characteristics of the motor imagery EEG, the EMD method is proposed to decompose the EEG signal into set of stationary time series called *Intrinsic Mode Functions* (*IMF*). These *IMF*s are analyzed with the bandpower (*BP*) to detect the caracteristics of sensorimotor rhythms (*mu* and *beta*). Finally, the data were reconstructed with only with the specific *IMF*s and then the band power is employed on the new database. Once the new feature vector is reconstructed, the classification of motor imagery is applied using *Hidden Markov Models* (*HMM*s). The results obtained show that the *EMD* method allows the most reliable features to be extracted from *EEG* and that the classification rate obtained is higher and better than using only the direct *BP* approach. Such a system appears as a particularly promising communication channel for people suffering from severe paralysis, for instance for persons with myopatic diseases or muscular dystrophy (MD) to move a joystick to a desired direction corresponding to the specific motor imagery.**

I. Introduction

Brain Computer Interfaces (*BCI*) is a direct communication pathway between a brain and an external device. The major goal of the BCI research is to develop systems which help disabled users to communicate with other people in order to control their artificial limbs or environment [1], [2]. A *BCI* system is represented as a system in a continuous closed loop, generally composed of six steps, (Figure 1): 1. Brain activity measurement, 2. Preprocessing, 3. Feature Extraction, 4. Classification, 5. Translation into a command and Feedback.

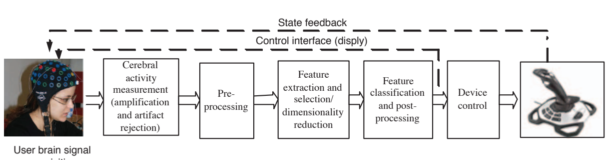

Fig. 1. General architecture of an online (*BCI*)

One major challenge of *BCI* is to describe the signals *EEG* by a few relevant values called features i.e. step 3. These features can then be used in step 4 in order to classify the users mental state. Many Physiological studies [4], [5], showed that brain rhythms related to motor actions are called sensorimotor rhythms, which are mainly located in the mu (8-12 Hz) and beta (13-30 Hz) frequency bands. A common approach in BCI is thus to extract the band powers of the rhythms mu and beta from the EEG signal and use them as classification features. Several common band power techniques were employed in the BCI literature. Herman et al. [6] demonstrated that the Yule and Welch PSD approaches, mainly dominate the other studied ones. These approaches are based essentially on some linearity and stationarity hypothesis such as the use of fast Fourier transform *(FFT)* spectrum in short time of a segment of data. The accuracy of the *FFT* calculation is closely related to the choice of the duration of the signal segment [7]. However, the nature of the EEG signal is nonstationary and nonlinear [8]. The main nonstationary and nonlinear feature extraction technique is the*Wavelet Transform* (*WT*) [9]. Despite being more effective than the *FFT*, this approach shows much bigger ambiguity in signal decomposition. However, it cannot provide higher resolution both in time and frequency domain, besides, the decomposition of signal is not adaptive. In this paper, we applied a recent technique proposed by Huang et al, called the*empirical mode decomposition* (*EMD*) for nonlinear and nonstationary time series data for pattern extraction from motor imagery EEG of left and right hand imaginary movement. This method *EMD* proposed by Huang et al. [10] is a data driven approach (i.e. one does not need to define a mother wavelet beforehand) that can be used to decompose adaptively a signal into a finite number of mono-component signals, which are known as *intrinsic mode functions* (*IMF*s) or *modes*. It considers signals at their local oscillations, but they are not necessarily considered in the sense of Fourier harmonics. Their extraction is non-linear, but their recombination for exact reconstruction of the signal is linear. The IMFs admit well-behaved *Hilbert transforms*

(*HT*) [11] and they satisfy the following properties: they are symmetric, different *IMFs* yield different instantaneous local frequencies as functions of time that give sharp identifications of embedded structures.

In this work, we propose an hybrid approach combining the *EMD* and *BP* for feature extraction from the *EEG* signals. We first apply the *EMD* to select only the *IMFs* corresponding to sensorimotor rhythms (*mu* and *beta*) using Welch-based *PSD* to extract the reliable information of *EEG* during left and right hand movement imagination.

Based on these new features, some valuable results were achieved by using a nonlinear classifier known as *hidden Markov models* (*HMMs*) [16].

II. Data bases and methods

A. EEG Data

The raw motor imagery EEG data is recorded in our reaserch lab from ten healthy non-experinced volunteers. The experiment is described as follow: ech subject was seated in an armchair and looked at a computer monitor. He was asked to keep their arms and hands relaxed and to avoid eye movements during the recording. Each trial (8s) began with a blank screen. After two seconds a warning stimulus was given in form of a beep. From 3 to 4.25s, an arrow (cue stimulus), pointing to the left or right, was shown on the screen. The subject was instructed to imagine a left or right hand movement until the end of the trial, depending on the direction of the arrow (Fig. 2). The experiment data were sampled at 256 Hz and filtered between 0,5 and 30Hz, then a notch filter was used to suppress the 50Hz power line interference. Two bipolar recordings overlying the left and right sensorimotor area were obtained by two electrodes C3 and C4 placed according to the international 10/20 system [19].

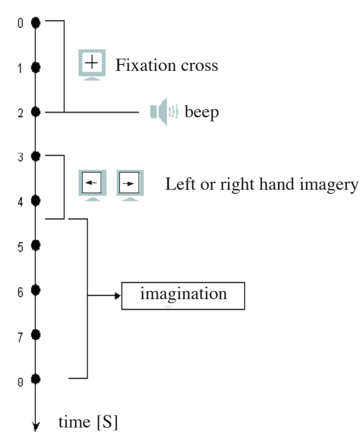

Fig. 2. Paradigm: Timing of one trial in the experiment.

B. Feature Extraction

1) *Empirical mode decomposition* (*EMD*) The traditional *EMD* was recently proposed [10] as an adaptive time-frequency data analysis method. It is defined by an algorithm based on an empirical framework. In most cases, the studies (performance, analysis,...) carried out on the *EMD* are done with extensive digital simulations in controlled conditions [10]. Despite the lack of theoretical formalism, this algorithm showed its capacity to analyze the signals. Using a new formulation for *EMD* based on constrained optimization, the results of [20] were very similar to those obtained with the traditional *EMD* algorithm.
The basic *EMD* is defined by a process called *sifting* to break down any multimodal signal to a sum of basis components called *intrinsic mode functions* (*IMFs*). The *IMFs* are zero-mean AM-FM signals which must satisfy two conditions: the first one is that the number of extrema and that of zero-crossing must differ at most by one; the second one is that the mean value between the upper and lower envelopes are equal to zero at any point. Conceptually, the establishment of this method is quite simple: one needs to consider a signal at its local oscillation level, remove the fastest oscillation and iterate the process on the residue considered as a new signal. At the end of the sifting processes, a given signal $x(t)$ can be written as a sum of a finite number of *IMFs*, $I_m(t), m = 1,\ 2, ..., M$, and a final residue $r_M(t)$:
$$x(t) = \sum_{m=1}^{M} I_m(t) + r_M(t).$$
The decomposition is stopped at step M, if either the residue $r_M(t)$ is a mono-component signal or has only 2 extrema [10]. The *stopping criterion* must be set to ensure that the obtained signal satisfies the properties of an *IMF* while limiting the number of iterations. For more details about the different steps of the *sifting* process for the calculation of the IMF_i as well as the *stopping criterion* definition see [10]. Since the decomposition into *IMFs* is based on the local characteristic time scale of the data, it applies to nonlinear and non-stationary processes.
2) *Band powers* (*BP*) The features may be extracted from the *EEG* signals by estimating the power distribution of the *EEG* in predefined frequency bands. In general, the band power is estimated by digitally bandpass filtering the data, squaring and averaging over consecutive samples according to a given window size. Pfurtscheller et al. [21] used the *BP* and demonstrated that for each subject, different frequency components in the mu and beta bands were found which provided best discrimination between left and right hand movement imagination. These frequency bands varied between 9 and 14 Hz and between 18 and 26 Hz.
3) Theory and application of hybrid method (*EMD-BP*) for motor for motor imagery: In this work, we propose a direct nonlinear approach to extract the more relevant *IMF*s corresponding to the different frequency when a subject imagines a movement of the right or the left

hand. Then, each *IMF* is analyzed by the Welch-based *PSD* to find the active frequency bands μ and β. Once the rhythms are found, the signal *EEG* will be reconstructed with only the *IMF*s that contain the frequencies corresponding to motor imagery. Fig. 3 summarizes the proposed feature extraction approach. The feature vector $\mathbf{p}_i$ used for the demonstration in this paper is composed, for each sample $i, 1 \leq i \leq 2048$, in a given trial (among a total of 160 trials) (see section II-A).

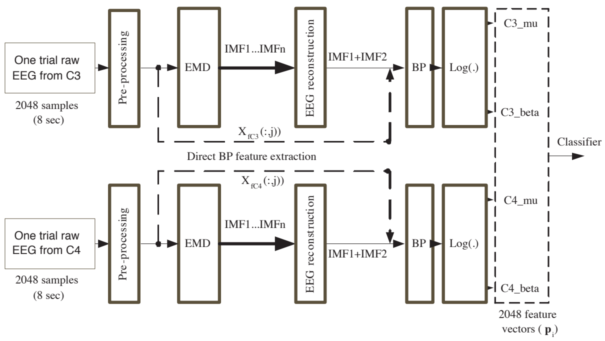

Fig. 3. Hybrid EMD-BP approach for one trial feature extraction.

We applied the *EMD* method on the *EEG* data defined in section II-A. Fig. 4 shows the result of one-trial (left hand movement imagination) EMD decomposition for subject 2 in the channels C3 and C4 respectively. Each channel is decomposed into ten *IMFs* and one residue.

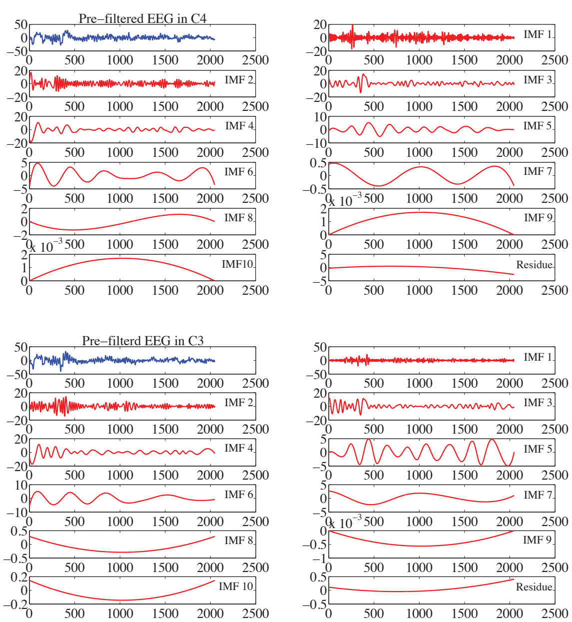

Fig. 4. The EMD decomposition results for subject 2 when he imagines left hand movement. From top-left to down-right: the raw signal, the ten *IMFs* and the residue in channel C4 (a) and in C3 (b)

To find out the characteristic IMF of μ and β rhythms, Welch based PSD method was adopted. So we calculated the mean power spectral density of all trials on the corresponding IMFs of C3 and C4. In Fig 5 we plotted the variation of the *PSD* versus frequency of each *IMF*.

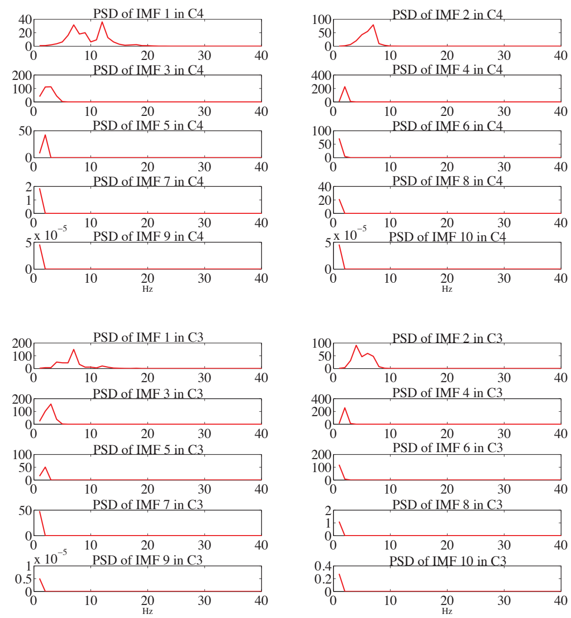

Fig. 5. The EMD decomposition results for subject 2 when he imagines left hand movement. From top-left to down-right: the raw signal, the ten *IMFs* and the residue in channel C4 (a) and in C3 (b)

We can notice that the characteristics of the active frequency bands corresponding to mu $[8-12Hz]$ and beta $[13-30Hz]$ are located only in *IMF*1, *IMF*2 on C3 and C4. Therefore, the new signal is reconstructed by keeping only the two first *IMF*s. *EMD* also allows to eliminate the artifacts in the *EEG* during the recording sessions like eye blinks and eyeball movements. In Fig 5, we noted that ocular artifact frequency is generally low around 5Hz with a high amplitude. This artifact appears mainly in *IMF3* and *IMF4*. Finally, band power was applied for the new signal.

C. *Classification*

1) Hidden Markov Models HMM *:* This method is very efficient nonlinear technique used for the classification of time series [16]. it necessitates two stages: a training stage where the stochastic process models are estimated through extensive observation corpus and decoding or detection stage where the model may be used off/on-line to obtain the likelihoods of the given test sequence evaluated by each model [24], [25]. A *HMM* is defined by the following compact notation to indicate the complete parameter set of the model $\lambda = (\mathbf{\Pi}, \mathbf{A}, \mathbf{B})$, where $\mathbf{\Pi}$, $\mathbf{A}$ and $\mathbf{B}$ are the initial state distribution vector, matrix of state transition probabilities and the set of the observation probability distribution in each state, respectively [16]. This set of parameters is defined by

$$\mathbf{\Pi} = [\pi_1, \pi_2, \ ..., \pi_N], \ \pi_i = P(q_1 = s_i),$$

$$\mathbf{A} = [a_{ij}],\ a_{ij} = P(q_{t+1} = s_j | q_t = s_i).$$

Where $1 \leq i, j \leq N,\ s_i,\ s_j \in S,\ S = \{s_1, s_2,\ ..., s_N\}, t \in \{1, 2,\ ..., T\}$. The observation at time (or index) t, $\mathbf{O}_t$, is considered in this paper as continuous $\mathbf{O}_t \in \mathbb{R}^K$. For a continuous observation, the state conditional probability of the observation $b_i(\mathbf{O}_t)$ may be defined by a finite mixture of any log-concave or elliptically symmetric probability density function (pdf), e.g. Gaussian pdf, with state conditional observation mean vector $\boldsymbol{\mu}_i$ and state conditional observation covariance matrix $\boldsymbol{\Sigma}_i$. In this paper we consider only a single Gaussian pdf, so $\mathbf{B}$ may be defined as $\mathbf{B} = \{\boldsymbol{\mu}_i,\ \boldsymbol{\Sigma}_i\},\ i = 1, 2,\ ..., N$. At each instant of time t, the model is in one of the states $i,\ 1 \leq i \leq N$. It outputs $\mathbf{O}_t$ according to a density function $b_i(\mathbf{O}_t)$ and then jumps to state $j,\ 1 \leq j \leq N$ with probability a_{ij}. The state transition matrix defines the structure of the *HMM* [16]. The model λ may be obtained off-line by a given training procedure. In practice, given the observation sequence $O = \{\mathbf{O}_1, \mathbf{O}_2,\ ..., \mathbf{O}_T\}$, and a model λ, the *HMMs* need three fundamental problems to be solved:

1) How to calculate the likelihood $P(O|\lambda)$? The solution to this problem provides a score of how O belongs to λ.
2) How to determine the most likely state sequence that corresponds to O? The solution to this problem provides the sequence of the hidden states corresponding to the given observation sequence O.
3) How to adjust the model λ in order to maximize $P(O|\lambda)$? This is the problem of estimating the model parameters given a corpus of training observations sequences.

Problems 1 and 2 are solved in the decoding or detection stage using the forward or the Viterbi algorithms [16], while problem 3 is solved during the training phase using either a conventional algorithm such as the Baum-Welch algorithm [16].

Our training scheme is based on Baum-Welch algorithm and the *Bayesian Inference Criterion* (*BIC*) [26] [27]. This scheme makes the training procedure independent of the initialization problem and the a priori knowledge of the number of states in each *HMM* needed in the Baum-Welch training algorithm.

III. Results and Discussion

A. *Confusion matrix* HMM

The recognition rates calculated by *HMM* for the 10 subjects is represented by the confusion matrix (Table I) for the two feature extraction methods direct *BP* and *EMD + BP*. For each subject, the EEG data contains 160 trials and each trial lasts 8 seconds as shown in Fig 2 (a set of 80 trials for left hand movement imaginations and a set of 80 trials for right hand movement imaginations). Each set of movement imagination data was divided into two subsets for each mental task movement (40 trials for HMMs training and 40 trials for test). For the HMMs training and test steps, each trial (C3 or C4 feature sequence) is composed of T = 2048 samples, where each sample is a scalar value (K=1). In these subsets, we considered only the imagination period: 4 seconds to 8 seconds. We trained one HMM for each subject and for each motor imagery (right hand or left hand movement) using the data corresponding to that subject and that mental task. The number of states in each *HMM* was determined automatically by the *BIC*. This number is 2 in each *HMM*. These results show the interest of using the *EMD* method. It is clearly seen that the combination of the two features methods, *EMD* and *BP* gives the best classification rates.

TABLE I
Confusion matrix: recognition rates by *HMM* classifier for Right movement imagination *RMI* and Left movement imagination *LMI*; using both methods of feature extraction *FE* for all subjects.

Subject	FE	true class	predicted class RMI	predicted class LMI
1	BP	RMI	**70**	30
		LMI	2.5	**97.5**
	EMD+BP	RMI	**90**	10
		LMI	2.5	**97.5**
2	BP	RMI	**77.5**	22.5
		LMI	7.5	**92.5**
	EMD+BP	RMI	**75**	25
		LMI	2.5	**97.5**
3	BP	RMI	**57.5**	42.5
		LMI	35	**65**
	EMD+BP	RMI	**72.5**	27.5
		LMI	40	**60**
4	BP	RMI	**71.42**	28.58
		LMI	37.15	**62.85**
	EMD+BP	RMI	**71.42**	28.58
		LMI	31.43	**68.57**
5	BP	RMI	**52.5**	47.5
		LMI	42.5	**57.5**
	EMD+BP	RMI	**72.5**	27.5
		LMI	40	**60**
6	BP	RMI	**80**	20
		LMI	50	**50**
	EMD+BP	RMI	**85**	15
		LMI	47.5	**52.5**
7	BP	RMI	**80**	20
		LMI	35	**65**
	EMD+BP	RMI	**80**	20
		LMI	32.5	**67.5**
8	BP	RMI	**87.5**	12.5
		LMI	32.5	**67.5**
	EMD+BP	RMI	**90**	10
		LMI	37.5	**62.5**
9	BP	RMI	**60**	40
		LMI	45	**55**
	EMD+BP	RMI	**82.5**	17.5
		LMI	25	**75**
10	BP	RMI	**60**	40
		LMI	42.5	**57.5**
	EMD+BP	RMI	**75**	25
		LMI	35	**65**

B. *Cohen's Kappa coefficient*

The evaluation of the performance of our feature extraction is based on the performance of the classification results on *RMI* and *LMI*. To quantify the level of performance achieved, we used the Cohen's kappa coefficient [29] to measure the agreement between the results of *HMM* classifier. The *Cohen's*

κ, is used to measure interobserver variability, i.e., how often 2 or more observers agree in their interpretations. The *Cohen's kappa* κ is defined by:

$$kappa = \frac{P_o - P_e}{1 - P_e}.$$

With P_o the *observed agreement* and P_e the *expected agreement*. Table II gives the scores and their averages for each one of the ten subjects in order to evaluate our hybrid method for feature extraction *EMD+BP*. Therefore, the use of *EMD+BP* approach gives to the *k scores* a significant superiority compared to the direct *BP* approach.

TABLE II
K VALUES OBTAINED BY *HMM* FOR *BP* AND *BP-EMD* FEATURE EXTRACTION METHOD. THESE VALUES CALCULATED FOR THE 10 SUBJECTS DURING *RMI* AND *LMI* AND THEIR MEAN VALUES.

Subjetct	*BP*	*EMD+BP*
1	0.6	0.8
2	0.8	0.8
3	0.4	0.6
4	0.2	0.4
5	0.2	0.4
6	0.4	0.4
7	0.4	0.4
8	0.6	0.6
9	0.2	0.6
10	0.2	0.4
Mean *K*	**0.40**	**0.54**

C. Translation into a command

Once the motor imagery is identified, a command may be associated to this mental task in order to control a machine such as a hand prosthesis (opening/closing) [30]or a robot [31]. In this work, we constructed a new Simulink/MathWork model to translate on-line the EEG signals into low-level commands. Fig. 6 shows our experimental EEG-based BCI system.

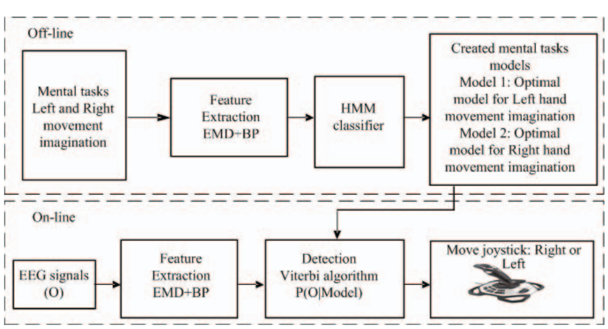

Fig. 6. the general conception of our asynchronous system BCI (offline - online) for reinforcement of a joystick movement.

1) **Off-line phase:** In the first step, the EEG signals are recorded while subjects imagine a right and left hand movement. The second step is the preprocessing and feature extraction of *EEG* data. In this step, we implemented our method to extract the relevant features of the *EEG*. This method is based on the combination of EMD and BP. In the third step, we implemented the *HMM* classifier to assign a model to each motor imagery task: λ1for LMI and λ2 for RMI.
2) **On-line phase:** Once the motor imagery is identified by one of the two models, a low-level command can be then associated to this mental task. This mechanism was implemented with Simulink/MathWork. EMD + BP and Viterbi algorithm [16] are implemented as an embedded function in Simulink in order to identify the motor imagery on-line. Viterbi algorithm has two inputs data, the first input is the *EEG* data and the second is the two models already constructed in the first step (offline). The Viterbi-based recognition result is translated into a command to reinforce the movement of the joystick (right or left) in order to help persons with myopathic diseases or muscular dystrophy to move this joystick to a desired direction.

IV. Conclusion

In summary, we have studied *EEG* signal processing, feature extraction and classification techniques in order to design *BCI* systems. We have first proposed contributions at the *EEG* signal processing to extract the most relavent information corresponding to motor imagery. We have proposed a combination of a nonlinear feature extraction technique based on the Empirical Mode Decomposition *(EMD)* and the bandpower *(BP)*. This method consists first in applying the (EMD) on the *EEG* signal to obtain the *IMF*s. Then, frequency analysis of these *IMF*s showed that the frequencies of the first *IMF*s correspond to the sensorimotor rhythms μ and β during both imaginary left and right hand movement. We also showed that the last *IMF*s correspond to artefacts and noise. Finally, the conventional *BP* of the active frequencies is applied to the relevant reconstructed signal. A nonlinear classifier based on Hidden Markov models *HMM* is employed to evaluate this feature extraction approach. Therefore, we deduce that the classification rate in the two movement imagination are better using *EMD+BP* approach than the direct *BP* approach. When subject motor imagery is recognized, we translated the *EEG* signal to a low-level command. This system allows subjects who suffer from severe motor disabilities to better reinforce a joysticks movement to right or left.

References

[1] J. R. Wolpaw, N. Birbaumer, J. McFarland D., G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control", *Clinical Neurophys*, vol. 113, 2002, pp. 767-791.
[2] J. R. Wolpaw, "Brain-computer interfaces (BCI's) for communication and control: Current status," *2nd Int. BCI Workshop Training Course*, 2004, pp. 29-32.
[3] S. G. Mason and G.E. Birch, "A general framework for brain-computer interface design," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,*11(1):70-85, 2003.
[4] G. Pfurtscheller, EEG Event-related Desynchronisation (ERD) and Event-related Synchronisation (ERS). In Niedermeyer, pp. 958-967, E., Lopes da Silva, F.H. (Eds.), Electroencephalography. Basic Principles, Clinical Applications, and Related Fields, forth ed., Williams and Wilkins, Baltimore.

[5] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*," vol. 89, no. 7, pp. 1123-1134, 2001.

[6] P. Herman, G. Prasad, T. M. McGinnity, and D. Coyle, "Comparative analysis of spectral approaches to feature extraction for eeg-based motor imagery classification," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 16, no. 4, 2008, pp. 317-326.

[7] T. Penzel and R. Conradt, "Computer based sleep recording and analysis", Sleep Medicine Reviews, vol. 4, no. 2, pp. 131-148, 2000.

[8] M. Lo, P. Tsai, P. Lin, C. Lin, and Y. L. Hsin, "The Nonlinear and non-stationary Properties in EEG Signals: Probing the Complex Fluctuations by Hilbert-Huang Transform," presented at Advances in Adaptive Data Analysis, pp.461-482, 2009.

[9] S. G. Mallat, "A theory for multiresolution signal decomposition : the wavelet representation," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674-693, July, 1989.

[10] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. of the Royal Society, vol. 454, no. 1971, pp. 903-995, 1998.

[11] D. Gabor, "Theory of communication," Proc. IEE, vol. 93 (III), pp. 429-457, 1946.

[12] C. M. Sweeney-Reed and S. J. Nasuto, "A novel approach to the detection of synchronisation in EEG based on empirical mode decomposition", J Comput Neurosci vol. 23, pp. 79111, 2007.

[13] G. Xiaojing, W. Xiaopei, and Z. Dexiang, "Motor Imagery EEG Detection by Empirical Mode Decomposition," *Conf Proc IEEE IJCNN 2008*, pp. 2619-2622, 2008.

[14] L. Wang, X. Guizhi, J. Wang, S. Yang, and W. Yan, "Application of Hilbert-Huang transform for the study of motor imagery tasks," *Conf Proc IEEE Eng Med Biol Soc.*, pp. 3848-3851 , 2008.

[15] R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern Recognition," 2nd. ed. (New York: Wiley-Interscience), 2001.

[16] L. R. Rabiner, "A tutorial on hidden Markov models and selected application in speech recognition", *Proc. IEEE*, vol. 77, no. 2, February 1989, pp. 267-296.

[17] Guger Technologies, "BSanalyze advanced biosignal processing and analysis version 3.07a,". *http://www.gtec.at/*.

[18] *http://www.gtec.at/*

[19] E. Niedermeyer, F. Lopes da Silva, "Electroencephalography: Basic Principles, Clinical Applications, and Related Fields," Lippincott Williams & Wilkins, 2004 ISBN 0781751268, 9780781751261.

[20] S. Meignen and V. Perrier, "A New Formulation for Empirical Mode Decomposition Based on Constrained Optimization," Signal Processing Letters, IEEE, vol. 14, pp. 932-35, 2007.

[21] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, "EEG-based discrimination between imagination of right and left hand movement," Electroenceph clin Neurophysiol, vol. 103, pp. 642-651, 1997.

[22] P. D. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms," IEEE Transactions on Audio Electroacoustics, vol. AU-15, pp. 7073, 1967.

[23] C. M. Bishop, "Neural Networks for pattern Recognition," Clarendon Press,Oxford,1995.

[24] T. Al-ani and D. Trad, "Signal Processing and Classification Approaches for Brain-computer Interface", In : Intelligent and Biosensors, Edited by Vernon S. Somerset ISBN 978-953-7619-58-9, pp. 25-66, Publishing date: January 2010.

[25] D. Trad, S. Helmy, T. Al-ani, and S. Delaplace, "Hidden Markov Models and Support Vector Machines applied to P300 Based Brain-Computer Interface - a Comparative study", ICTA 09, Hammamet, 7-9 Mai, 2009.

[26] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, Vol. 6, no. 2, pp. 461-464, 1978.

[27] S. Helmy, T. Al-ani, Y. Hamam, and E. El-madbouly, "P300 Based Brain-Computer Interface Using Hidden Markov Models," ISSNIP 2008, Sydney, 15-18 Dec., 2008.

[28] T. Al-ani, "Hidden Markov models in dynamic system modelling and diagnosis, Hidden Markov Models, Theory and Applications," ISBN: 978-953-307-208-1, InTech, In Press.

[29] J.Cohen,"A Coefficient of Agreement for Nominal Scales", Educational and Psychological Measurement, 20(1) :3746, April 1960.

[30] C. Guger, W.Harkam, C.Hertnaes. and G.Pfurtscheller, "Prosthetic control by an EEG-based brain-computer interface (BCI)", In Proceedings AAATE99, pp. 590-595.1999.

[31] A.Kbler, V.K.Mushahwar, L.R.Hochberg, and J.P.Donoghue." BCI meeting 2005-workshop on clinical issues and applications". In IEEE Transactions on Neural Systems and Rehabilitation Engineering. Vol. 14, No. 2, pp. 131-134, 2006.