Proceedings of the
7th Graz Brain-Computer Interface Conference 2017 DOI: 10.3217/978-3-85125-533-1-11

INCREASING STROKE PATIENTS MOTOR IMAGERY CLASSIFICATION BY SELECTING FEATURES WITH PARTICLE SWARM OPTIMISATION

J. Cantillo-Negrete[1], R.I. Carino-Escobar[1], P. Carrillo-Mora[2], J. Gutierrez-Martinez[1]

[1] Division of Medical Engineering Research, Instituto Nacional de Rehabilitación, Mexico City, Mexico

[2] Division of Neurosciences, Instituto Nacional de Rehabilitación, Mexico City, Mexico

E-mail: jcantillo@inr.gob.mx

ABSTRACT: Motor Imagery based Brain-Computer Interfaces (BCI) have shown potential for the rehabilitation of stroke patients. In order to make BCI systems available in the clinical environment new processing stages that increase the number of patients that can use these systems must be developed. This work presents a novel processing stage for BCI systems using the Filter Bank Common Spatial Patterns algorithm for feature extraction and Particle Swarm Optimisation for feature selection. The proposed BCI's processing stage performance was evaluated with electroencephalography data of six stroke patients, which performed motor imagery of their paralysed hand. Offline tests reached average classification accuracies of 75±8 %. For 4 out of 6 patients, the proposed method showed a statistically significant higher performance ($p<0.05$) than the Common Spatial Pattern method. Therefore, although a higher sample is needed to confirm the observations, it is possible to significantly improve hand motor imagery classification by selecting filter bank common spatial patterns features with particle swarm optimization.

INTRODUCTION

Stroke is the first cause of disability worldwide [1]. Approximately 400 patients receive neurorehabilitation therapy for stroke sequelae each year in the National Institute of Rehabilitation, located in Mexico City. Loss of motor function (known as hemiparesis) is one of the most disabling consequences of stroke, which usually affects both upper and lower limbs from one side of the body.

Assistive technologies such as Brain-Computer Interfaces (BCI) provide an artificial communication channel between the brain and an external device such as a robotic orthosis [2, 3]. BCI systems based on motor imagery (MI) of affected limbs have shown great potential as a tool for brain plasticity enhancement [4, 5]. MI is a mental rehearsal of movements of a limb, for example the hand or foot, without muscle activation [6, 7, 8]. MI elicits distinctive patterns in the electrical activity of the sensory-motor cortex, mainly in the frequency bands known as mu (8-13 Hz) and beta (14-30 Hz) [6, 9]. A MI-based BCI system is comprised of four stages: acquisition, pre-processing, feature extraction and classification. Most BCI acquire electroencephalography (EEG) since is a non-invasive technique, has a good time resolution and is easy to accept by patients. Linear Discriminant Analysis (LDA) is the most used classification technique reported in BCI publications [10, 11]. One of the most effective feature extraction methods is the Common Spatial Patterns (CSP) algorithm, which computes a set of spatial filters that optimally differentiate two classes of MI. To achieve good classification performances using the CSP algorithm, the temporal filtering of the EEG signal must be performed on a specific frequency band, usually this band is comprised by the mu and beta frequency range. Two other parameters that need to be set up are the time interval from which features are going to be extracted, and the subset of spatial filters involved in the feature extraction process [12].

The performance of CSP can be enhanced by selecting subject-specific parameters. Therefore, modifications to the original CSP method have been proposed to include this aspect. One of such modifications is known as Filter Bank Common Spatial Patterns (FBCSP); this method performs an automatic frequency band selection for temporal filtering of the EEG [13]. FBCSP algorithm employs a filter bank that decomposes the EEG into 9 different frequency bands covering the range of 4 to 40 Hz. Each of these 9 frequency bands is spatially filtered using the CSP algorithm; afterwards the extracted features for each band are selected with either the Mutual Information-based Best Individual Feature (MIBIF) or the Mutual Information-based Rough Set Reduction (MIRSR) algorithms. Classification is performed only with the selected features [13,14]. Feature selection is an important stage of the FBCSP algorithm, since it lowers the number of frequency bands needed for MI classification, and at the same time increases the classification performance of the BCI system. Feature selection is in fact an optimisation problem, and therefore artificial intelligence techniques, such as Particle Swarm Optimisation (PSO), could be used for finding a solution for it. PSO was originally proposed by Shi and Eberhart, inspired by the social behaviour of bird flocks

 DOI: 10.3217/978-3-85125-533-1-11

while searching for food. PSO performs a search in the space of the problem, with the aid of a population (called swarm) of individuals (called particles). Each particle executes a search based on its current position and velocity in the search space. In each iteration (called generations), the position and velocity of the particles are updated according to their best previous position (local search) and the best position of the swarm (global search) [15]. To the author's knowledge, there are few studies that describe the use of PSO as a feature selection algorithm for BCI systems [16,17].
In this work, a novel signal processing stage comprised of FBCSP for feature extraction, PSO for feature selection and LDA for classification was implemented as part of a BCI system. The proposed algorithm was evaluated offline with data of patients with subcortical stroke diagnosis.

MATERIALS AND METHODS

Participants: The sample for this study comprised 6 patients diagnosed with stroke (Mean = 55.8 ± 12 years). In order to be considered for inclusion in the study, patients had to have a first stroke event of subcortical localisation, confirmed by a neurologist by means of neuroimaging studies (Magnetic Resonance or Computed Tomography); total or partial paresis of one of their hands; without clinical history of any other previous neurological or psychiatric diseases; right handed; with normal or corrected to normal vision and, with a normal performance in the subscales of digit detection and visual detection of the neuropsychological test NEUROPSI (this test has been validated for Spanish-speaking populations) [18]. The subscales evaluate the ability to follow instructions and concentrate in repetitive tasks. Subcortical stroke patients were selected since their brain damage does not involve the brain cortex and, therefore, they were less likely to present significant cognition impairments. Patients' data are shown in Tab. 1.

Table 1: Clinical and Demographic data of patients

Patient	Age	Gender	Hemiparesis	Evolution
1	50	Male	Right	7 months
2	57	Female	Right	36 months
3	58	Male	Left	2 months
4	79	Female	Left	1 month
5	46	Male	Left	3 months
6	45	Male	Left	3 months

EEG acquisition: A g.USBamp biosignal amplifier from g.tec was used for EEG acquisition. EEG was acquired with 24-bits of resolution and sampling rate of 256 Hz. Active EEG electrodes were used for acquisition, with 11 electrodes placed over the scalp of the patients, in positions C3, C4, Cz, T3, T4, F3, F4, Fz, P3, P4 and Pz of the international 10-20 system. Ground placement was set in the AFz position, and the reference electrode was placed in the right earlobe. To verify that no real movements were elicited during MI, Electromyography (EMG) was recorded from the deep flexor and superficial muscles of the fingers of both hands. For each patient, four recording sessions were performed in consecutive days, with 120 trials recorded in total. Recordings were performed in 4 days to avoid patients' exhaustion, and all trials recorded per patient were included in the analysis. Patients were instructed to sit in a comfortable armchair, with a computer monitor placed at 150 cm in front of them. Visual cues shown in the monitor directed the patients to perform both rest with eyes open and MI from their paralysed hand. EEG acquisition was performed using a similar strategy as the one followed by the Graz paradigm [19]. Fig. 1 shows that the rest interval of the trials lasted 3 s and the MI interval lasted 5 s.

Implementation of the FBCSP+PSO algorithm: A one-second length window was extracted from 1.5 s to 2.5 s to obtain the rest information for each trial. Another window of one-second length was extracted from the 3.5 to 4.5 s time interval of each trial, to obtain the MI information of the trials, as observed in Fig. 1. These time windows were selected based on previous studies which show that differentiation between MI and REST classes is higher in these time intervals [20]. The FBCSP algorithm encompassed the processing stage of the BCI system, and PSO was used for feature selection (named FBCSP+PSO). A diagram of the algorithm's implementation is shown in Fig. 2.

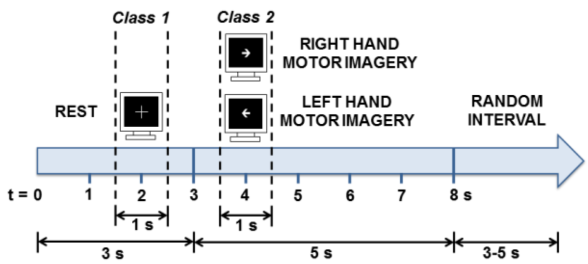

Figure 1: Illustration of the experimental paradigm. Dotted lines show the time windows extracted from EEG signals.

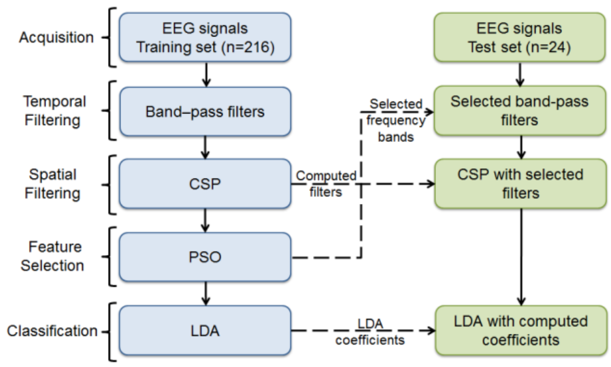

Figure 2: Diagram of FBCSP+PSO implementation

 DOI: 10.3217/978-3-85125-533-1-11

EEG data were filtered in order to obtain 6 frequency sub-bands, each 4 Hz broad, and with 1 Hz of overlapping in order to avoid loss of information. Encompassing both alpha and beta frequency bands as follows: 8-12 Hz, 12-16 Hz, 16-20 Hz, 20-24 Hz, 24-28 Hz and 28-32 Hz. A 60 Hz band-stop filter was also applied to the EEG signals. All filters were FIR filters of 20th order, selected for their linear phase features.
For the EEG data filtered in each sub-band, spatial filters were computed with the CSP algorithm. CSP performs a linear transformation on the EEG data, in order to obtain features whose variances are optimal for classification of two classes of MI, in a specific frequency band. Details of the CSP implementation can be found in the works of Blankertz et al. [21], and Ramoser et al. [22]. Spatial filters were computed using the MATLAB command $W = eig\,(S1, S1 + S2)$ as suggested in the above-mentioned works. W is the matrix containing the spatial filters, $S1$ and $S2$ are the covariance matrices of MI and rest computed from the EEG data of each filtered frequency sub-band. In the implementation of the original CSP, only the first and last m columns of the W matrix (m is generally 2) are used to generate the feature vector used for classification. With the goal of having a greater chance of finding the optimal sub-band for each patient, in this work all possible features were extracted with CSP. The feature vector generated for each trial i is comprised as follows:

$$f_i = [f_{1,i}, f_{2,i}, f_{3,i}, f_{4,i}, f_{5,i}, f_{6,i}] \qquad (1)$$

Therefore, CSP features computed for the training set comprised by nt trials are:

$$F_{Train} = [f_1; f_2; f_3; f_4; \dots; f_{nt}], \qquad F_{Train} \in \mathbb{R}^{nt\times 66} \qquad (2)$$

Where 66 are the 6 frequency band features f extracted for each of the 11 recorded electrodes. For feature selection, PSO was used for selecting a subset of features from F_{Train} in order to decrease both the classification error and the number of selected features. PSO was computed by solving two equations:

$$v_i^{n+1} = w \cdot v_i^n + c_1 \cdot r_1 \cdot (PBest_i^n - x_i^n) + c_2 \cdot r_2 \cdot (GBest_g^n - x_i^n) \qquad (3)$$

$$x_i^{n+1} = x_i^n + v_i^{n+1} \qquad (4)$$

Where x_i^{n+1} and v_i^{n+1} are the position and velocity of the ith particle of the nth generation. For PSO implementation 50 particles and 50 generations were used. w is the inertial weight of PSO which linearly descends from 1 to 0 as generations of PSO are computed. c_1 and c_2 are positive constants set to 1. r_1 and r_2 have random values between 0 and 1, which coupled to c_1 and c_2 set the local and global search properties of PSO. $PBest_i^n$ is the best position reached by the ith particle in the nth generation. $GBest_g^n$ is the best position (g) reached by the entire swarm in the nth generation. The maximum position value that a particle could reach was 1 and the minimum was 0. Maximum speed of each particle was set to 1 and minimum speed to 0. In this work, the search space of PSO was $1xD$, where D equals 66, and was comprised of the 66 features that can be selected from the FBCSP algorithm. Each computed solution with PSO is a subset of the selected features. Solution values are in the range from 0 to 1. If the value of an element of the solution was higher or equal to 0.5, then the corresponding feature was selected. The original CSP algorithm states that selected features must be paired, so complementary features of the selected ones were also included, in case they were not originally selected by PSO. Selected features from the training set were used for designing an LDA classifier. PSO fitness value was computed with the following equation:

$$value = (err\times 2) + (nselec/66) \qquad (5)$$

Where err is the computed classification error from the training set. $nselec$ is the number of selected features. Variables err and $nselec/66$ have values ranging from 0 to 1. Both parameters err and $nselec/66$ are summed, so that PSO can perform a reduction of both classification error and the number of features used for classification. The value err is multiplied by 2, so that the optimization priority of PSO is the reduction of the classification error over the selection of a lower number of features. The stop criteria used for PSO was either achieving 0% of classification error, or 50 generations. Fig. 3 shows a block diagram depicting the implemented PSO algorithm.

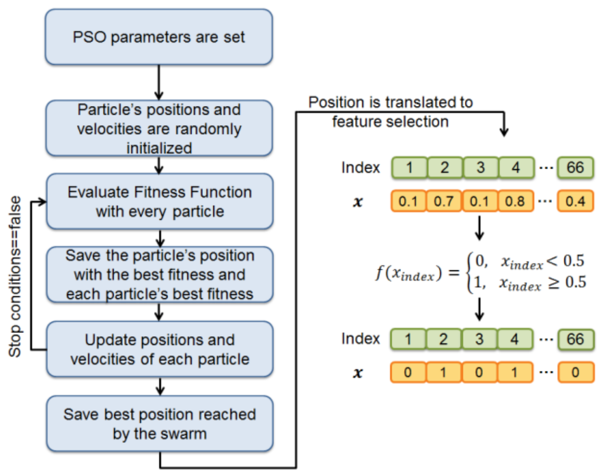

Figure 3: Block diagram describing the implementation of the PSO algorithm

With the final selected features (x) and the training set, a LDA classifier was designed, which was later evaluated with the testing set. Features selected with PSO in the training stage were the same as the ones used for the testing stage of the classifiers. LDA

 DOI: 10.3217/978-3-85125-533-1-11

performance was measured by computing the percentage of classification accuracy (%CA).

Cross-Validation: A stratified cross-validation of 10x10-Fold was used in order to avoid bias in the computation of %CA. Classifiers were tested using totally different datasets than the ones used for training. For each fold and repetition, the FBCSP+PSO algorithm was calculated. The 100 values of %CA obtained from this procedure were used to compute the average %CA for each patient.

For comparison purposes, the performance of the FBCSP+PSO method was compared with that of the original CSP using the same training and test subsets, and applied to a frequency band of 8 to 32 Hz.

Statistical Analysis: In order to assess the reliability of the BCI system, both %CA and the practical level of chance were computed. The practical level of chance for each experiment was not 50%, since its value needs to be computed by means of a confidence interval as explained by Muller-Putz et al. [23]. Practical level of chance was computed with a binomial distribution using a 95% confidence interval, with 120 trials encompassing the data of each class. The computed %CA were compared with the practical level of chance in order to assess if a patient could control the BCI system.

A paired t-test (α=0.05) was performed for comparing the %CA obtained with the proposed FBCSP+PSO method, and the original CSP (with a frequency band ranging from 8-32 Hz).

Computational cost: The averaged execution time of the proposed algorithm's training stage for each patient's cross validation was used to estimate its computational cost. All computations were performed in a PC with a 2.5GHz Core i7 processor and 12GB of RAM.

RESULTS

Tab. 2 shows the number of selected features by the FBCSP+PSO algorithm for each patient. This number is the mode from the 100 values computed from the 10x10-Fold cross-validation with the train set. On average, for each patient, 10 features were selected by PSO. The most selected frequency band for all experiment's repetitions is also shown: for 5 of the 6 patients it was from 8 to 12 Hz, which comprises the mu rhythm, while for the other patient the selected frequency sub-band was 12 to 16 Hz. Tab. 3 shows the %CA obtained with FBCSP+PSO and the ones obtained with CSP with a frequency band from 8 to 32 Hz are shown. These percentages are the offline MI and rest recognition capabilities of the BCI.

It is important to remember that the number of selected features with the CSP algorithm was always 4 ($2\times m$). An asterisk (*) was used to indicate if a statistically significant difference (p<0.05) was found between both methods. FBCSP+PSO showed better performance than CSP for the 6 patients. For 4 of the 6 patients, differences were statistically significant.

Table 2: Feature selection performed with PSO. SD refers to standard deviation.

Patient	FBCSP+PSO	
	Features	Frequency Band (Hz)
1	10	8-12
2	8	12-16
3	8	8-12
4	10	8-12
5	10	8-12
6	12	8-12
Mean(SD)	10(2)	-

Table 3: Performances of FBCSP+PSO and CSP. An asterisk (*) means that statistically significant differences (p<0.05) were found between both methods. SD refers to standard deviation.

Patient	FBCSP+PSO % Classification accuracy (SD)	CSP % Classification accuracy (SD)
1	83 (2)	82 (1)
2	85 (2)	84 (1)
3	68 (2)*	66 (1)
4	65 (3)*	58 (2)
5	76 (2)*	69 (1)
6	74 (2)*	63 (1)
Mean(SD)	75(8)	70(10)

The average computational cost of FBCSP-PSO training stage across all patients was 3.6 s (SD=0.04 s).

DISCUSSION

The presented novel processing stage was comprised by the FBCSP algorithm for feature extraction and PSO for feature selection. Test results were compared to those from the original CSP algorithm with a frequency band from 8 to 32 Hz. The proposed method was designed in order to increase the BCI's MI classification performance of the paralysed hand of stroke patients. Offline performances of the proposed processing algorithm achieved better performances than the original CSP. It is important to mention that for 4 out of 6 patients, these performance differences were statistically significant. These results are different from the ones presented by Ang et al., who performed an offline evaluation of the FBCSP that employed the MIRSR feature selection algorithm. They performed their test with a public database comprised of 9 healthy subjects. In their work, it is shown that FBCSP using the MIRSR algorithm had better performances for 6 out of 9 subjects than CSP (using a 7 to 35 Hz band), but none of the performance differences were statistically significant [14]. Therefore, the FBCSP+PSO method seems to be a better option for automatic frequency band selection of each patient.

The average offline performance computed for each patient is similar to the one reported by Ang et al. in a study which analysed the performance of 46 stroke

 DOI: 10.3217/978-3-85125-533-1-11

patients which achieved an average of 74% of correct classification. In order to acquire MI from the patients' paralysed hands, authors recorded 27 EEG channels. The processing stage comprised the FBCSP using MIBIF as feature selection algorithm [24]. In the present work, similar offline performances were obtained, however only 11 EEG channels were recorded. PSO is an optimisation method for which extensive research has been conducted in order to ensure better convergence and to reduce stagnation of the search space. The heuristic nature of PSO implies that the method performance will not be limited by statistical features of the search space, since the method does not need to compute inverse matrices or other computations which often present restrictions, especially for high dimensional search spaces. Consequently, PSO can be easily adapted for feature selection in MI-based BCI with setups involving a high number of EEG electrodes; however, one of the main disadvantages of PSO optimisation is the high computational cost required for its training phase. In this work, computational cost was not an issue since a relative low number of EEG channels were recorded and processed. Offline performances of the BCI system show that PSO implementation for feature selection of FBCSP allows this method to have better performances than CSP. This performance is achieved by setting a multi-objective optimisation for the PSO algorithm, which is computationally efficient since it only required computing the LDA performance and the number of selected features. It is important to mention that, in order to achieve better performances, higher importance was given to the LDA's classification performance than to the number of selected features in the fitness function.

One of the limitations of the present study was that scalp location of the selected features was not analysed. However, all the recorded electrodes were placed over the sensorimotor cortex and, therefore, in an online BCI aimed for neurorehabilitation no maladaptive changes during neural re-organization would be elicited by the feedback.

CONCLUSION

This work presents a novel processing stage for BCI systems. The proposed processing stage comprised of FBCSP+PSO combined with LDA showed good performances for classification of MI from the paralysed hand of stroke patients. PSO as a selection algorithm for FBCSP features allows reducing the problem's dimensionality and achieving better classification performances, compared to those obtained if only the original CSP is used. The next developing stage of the system will be to perform tests involving direct EEG acquisition from patients. An online implementation of the proposed algorithm must be assessed to further confirm its feasibility for stroke patients' rehabilitation.

ACKNOWLEDGEMENTS

The authors would like to thank the National Council of Science and Technology (CONACYT) from Mexico for funding under the grant number SALUD-2015-2-262061. The authors would also like to thank the patients for their participation in this study.

REFERENCES

[1] Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Cushman M, Ferranti S, et al. Heart Disease and Stroke Statistics-2015 Update A report from the American Heart Association. Circulation. 2015; 131: 29-322

[2] Pfurtscheller G, Guger C, Muller G, Krausz G, Neuper C. Brain oscillations control hand orthosis in a tetraplegic. Neurosci. Lett. 2000;292(3)

[3] King CE, Dave KR, Wang PT, Mizuta M, Reinkensmeyer DJ, Do AH, et al. Performance Assessment of a Brain-Computer Interface Driven Hand Orthosis. Ann Biomed Eng. 2014;40(10): 2095-2105

[4] Mihara M, Hattori N, Hatakenaka M, Yagura H, Kawano T, Hino T, et al. Near-infrared spectroscopy-mediated neurofeedback enhances efficacy of motor imagery-based training in poststroke victims: a pilor study. Stroke. 2013;44: 1091-1098

[5] Mrachacz-Kersting N, Jiang N, Stevenson AJT, Niazi IK, Kostic V, Pavlovic A, et al. Efficient neuroplasticity induction in chronic stroke patients by an associative brain-computer interface. J Neurophysiol. 2016;115: 1410-1421

[6] Pfurtscheller G, Silva L. Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin. Neurophysiol. 1999;110(11): 1842-1857

[7] Zimmermann-Schlatter A, Schuster C, Puhan M, Siekierka E, Steurer J. Efficacy of motor imagery in post-stroke rehabilitation: a systematic review. J Neuroeng Rehabil. 2008;5:8

[8] Pfurtscheller G, Brunner C, Schlogl A, Lopez da Silva FH. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. Neuroimage. 2006; 31(1): 153-159

[9] Andrew C, Pfurtscheller G. On the existence of different alpha band rhythms in the hand area of man. Neurosci Lett. 1997;222(2): 103-106

[10] Muller J, Pfurtscheller G, Flyvbjergc H. Designing optimal spatial filters for single trial EEG classification in a movement task. Clin Neurophysiol. 1999;110(5): 787-798.

[11] Ye J, Xiong T. Computational and Theoretical Analysis of Null Space and Orthogonal Linear Discriminant Analysis. J. Mach. Learn. Res. 2006;7: 1183-1204

[12] Blankertz B, Tomioka R, Lemm S, Kawanabe M, Müller KR. Optimizing spatial filters for robust EEG

 DOI: 10.3217/978-3-85125-533-1-11

single-trial analysis. IEEE Signal Proc Mag. 2008; 25:41–56.

[13] Ang KK, Chin ZY, Zhang H, Guan C. Filter Bank Common Spatial Pattern (FBCSP), in Brain-Computer Interface, IEEE IJCNN, Hong Kong, China, 2008, 2390-2397

[14] Ang KK, Chin ZY, Wang H, Guan C, Zhang H. Filter Bank Common Spatial Pattern algorithm on BCI competition IV Datasets 2a and 2b. Front. Nerosci. 2012; 6(39): 1-9

[15] Shi Y, Eberhart RC. A modified particle swarm optimizer, in proc. of the IEEE international Conference on Evolutionary Computation, Anchorage, United States, 1998, 69-73.

[16] Wei Q, Wei Z. Binary particle swarm optimization for frequency band selection in motor imagery based brain-computer interfaces. Bio-med Mater. Eng. 2015;26: 1523-1532

[17] Atyabi A, Luerssen MH, Powers D. PSO-based dimension reduction of EEG recordings: Implications for subject transfer in BCI. Neurocomputing. 2013;119: 319-331.

[18] Ostrosky-Solis F, Gómez-Pérez E, Ardila A, Rosselli M, Matute E, Pineda D. Batería Neuropsicolófica NEUROPSI Atención y memoria, 6 a 85 años de edad. Bookstore, Mexico.

[19] Pfurtscheller G, Neuper C. Motor imagery and direct brain-computer communication. P. IEEE. 2001;7: 1123:1134.

[20] Cantillo-Negrete J, Gutiérrez-Martínez J, Flores-Rodriguez TB, Cariño-Escobar RI, Elías-Viñas D. Characterization of electrical brain activity related to hand motor imafery in healthy subjects. Rev. Invest. Clin. 2014;66 Suppl 1: 111-121

[21] Blankertz B, Tomioka R, Lemm S, Kawanabe M, Müller KR. Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Proc. Mag. 2008;25: 41-56

[22] Ramoser H, Muller-Gerking J, Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE T Biom-med Eng. 2000;8(4): 441-446

[23] Müller-Putz JR, Scherer R, Brunner C, Leeb R, Pfurtscheller G. Better than random? A closer look on BCI results. Int. J. Bioelectromagn. 2008;10: 52-55

[24] Ang KK, Guan C, Chua KS, Ang BT, Kuah CW, Wang C. A Large Clinical Study on the Ability of Stroke Patients to Use an EEG-based Motor Imagery Brain-Computer Interface. Clin EEG Neurosci. 2011;42(4): 252-258