Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Engineering

Elixir Comp. Engg. 82 (2015) 32129-32133

ISSN: 2229-712X

Performance Analysis of PSO and GA Algorithms in Order to Classifying EEG Data

Masume Esmaeili[1], Morteza Zahedi[1] and Nasser Hafezi-Motlagh[2]

[1]Department of Computer Engineering& IT, Shahrood University of Technology, Shahrood, Iran.

[2]Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.

ARTICLE INFO

Article history:

Received: 24 November 2014;
Received in revised form:
19 April 2015;
Accepted: 30 April 2015;

Keywords

EEG signal,
Principle component analysis,
Genetic Algorithm,
Particle Swarm Optimization.

ABSTRACT

In this Research, a new method has been proposed in order to classify the mental tasks which represent the Electroencephalogram (EEG) signal as time series. Time series are kind of data format which depict signal voltage varieties in time domain. Different parts of the different signals have different powers, so in first step and in the preprocessing, signal partitioning into several fixed windows is needed. Toward the extracting appropriate features from each EEG signal window, PCA algorithm is used. So for each window, a feature vector is made by PCA, and a general vector is created from these primary vectors. In order to refuse redundancy caused by non-important windows, the best combination of such vectors, that have the best results in classification, should be probed. Toward this goal, two feature extraction methods, Particle Swarm Optimization (PSO) and Genetic Algorithm (GA), are applied. K-Nearest Neighbor (KNN) was used as fitness function for PSO and GA. These methods select such windows whose combination of feature vectors are best and increase TP (true positive) of the classifier. The results show that GA and PSO improve the power of classification, but GA is more efficient.

Introduction

A brain computer interface provides a direct pathway between the brain and an external device. It completely opens new communication channel without using any of peripheral nervous or muscles to help disabled people. In order to open this pathway, the brain's electrical activity, generated by neurons and postsynaptic potentials, must be monitored. Several invasive and non-invasive methods can be presented for recording brain activity. In invasive methods such as Electrocorticography (ECoG), single microelectrode (ME), microelectrode array (MEA), and local field potentials (LFPs), the skull should be split surgically and the electrodes should be placed inside the skull, on the cortex of the brain [1]. However, in non-invasive methods such as EEG recording, the brain activity is recorded by placing electrodes over the skull above the cortex [2], [3]. The main disadvantage of EEG recording and any other non-invasive methods compared to invasive ones, is their very weak spatial resolution which makes them hard to locate the exact spot of the activity. Because the skull causes spatial smearing of the signal, two third of any activity generated by the neurons is lost due to misalignment of the firing neurons. In fact, in this method any activity can only be measured on the surface of the cortex which leaves out the majority of the neurons, since the voltages measured are extremely low. But the reason why EEG recording was selected as the measurement method of brain activity is based on its ease of appliance: there is no need for splitting scalp to place electrodes. Other reasons are related to the portability of devices and excellent temporal resolution (milliseconds range). Any variation in brain activity will be registered almost instantaneously. An EEG recording based BCI system is designed to extract specific features of EEG activity and use them in order to control the system. Different algorithms are presented toward this goal. The main idea is the appropriate features extraction for appropriate classification.

In this paper, the proposed method uses PCA for extracting principle features from the EEG segments. Different partitions of each signal have different levels of power in classification, so the first step is partitioning signals into several windows and then extract fitting feature vectors from each window distinguishably. The final feature vector is made by placing these feature vectors side by side. To maximize the efficiency of the algorithm, two methods were used, PSO vs. GA. They test the power of the classifier with different combinations of windows for finding the best solution. Finally, the power of GA and PSO are compared. The results show that GA is a better method for finding the best solution in a big space.

Dataset of mental tasks

The dataset used in this paper is recorded by Researchers of CEBL laboratory of Colorado University. Seven chosen subjects were asked to perform five mental tasks. The subjects were seated in a sound controlled booth with dim lighting and noiseless fans for ventilation. In order to obtain the highest influence of EEG signal, the electric activity of different points of cerebral cortex had to be measured [4], [5]. One electrode cap elastic was used for recording EEG signals, from channels O_1, O_2, P_3, P_4, C_3 and C_4 in 10-20 system. The electrodes were connected through a bank of Grass 7P511 amplifiers and band pass filter with 0.1--100 Hz Range. Each task was recorded through 10 seconds and the sampling rate was 250 Hz with a Lab Master 12 bit A/D converter mounted in an IBM-AT computer. Each task was repeated five times per session [6].

Subjects were asked to perform the five mental tasks:

• **Baseline**: the subjects were asked to be relaxed with no thinking (about anything) and no movements.

• **Letter Task**: The subjects were shown images of words as each word was indicative of a friend or family member (e.g., “father”, “mother”, “aunt”, “uncle”, etc.); they were instructed to compose mentally a letter to a friend or relative without vocalizing or making any physical movements.

Tele:
E-mail addresses: esmaeili.m2011@gmail.com

• **Math Task**: The subjects were given nontrivial multiplication problems, such as 63 times 84, and were asked to solve them without vocalizing or making any physical movements.

• **Geometric Figure Rotation**: The subjects were shown images of three-dimensional figures, and asked to visualize them being rotated around an axis. The figures were all three-dimensional extrusions of randomly drawn two-dimensional shapes.

• **Visual Counting task**: the subjects were asked to imagine a blackboard and visualize numbers written on the board, one after another, sequentially in ascending order, as the previous numeral was to be erased before the next was written.

Proposed method

The Careful analysis of the EEG signals can provide valuable insight and improve understanding of the mechanisms detection. As EEG signals are non-stationary, those methods that analyze the signals in frequency domain are not highly successful. However, using techniques in time component domain or time-frequency domain in order to extracting efficient features, can provide appropriate results [6], [7]. The proposed method uses PCA for extracting features from EEG time Series. Different sections of the signals have different resolutions in classification. Therefore, the first step is partitioning signals into several windows, and then finding the best windows in classification by PSO and GA algorithms. For making the final feature vector by PSO or GA, KNN is used for classifying signals. Fig.1 shows the proposed method flowchart.

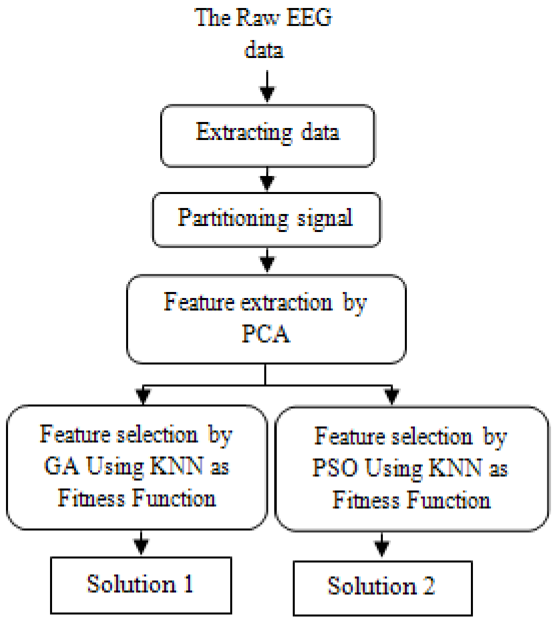

Fig 1. The Flowchart of Proposed Method

Partitioning signals

Before classifying EEG signals, the first step is preprocessing. Different sections of each signal have different power in classification signals. Therefore, a good way is dividing signals into several windows, and then extracting features window by window. In this paper, signals with 10-second length are divided into twenty windows with 0.5 second length. For extracting appropriate features from each window, PCA algorithm is used to make twenty feature vectors. From placing these vectors side by side, the final feature vector is created. Fig.2 shows the five first half-second windows of baseline task related to subject1. With removing some windows from signals, the performance of the classifier is increased. Thus, for obtaining the best structure of signals, first the power of each window in classification should be evaluated So, for each window of signals, we should classify test signals to evaluate that window. The success rate (TP) of each window is shown in Fig.3. Fig.3 shows that the highest TP achievement is in windows 8 and 10.

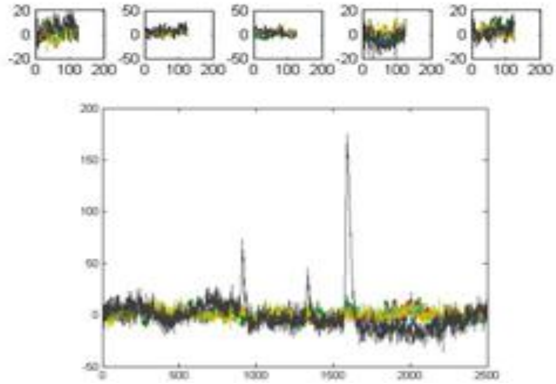

Fig 2. The EEG signal of baseline task related to subject1 and its five first half-second windows

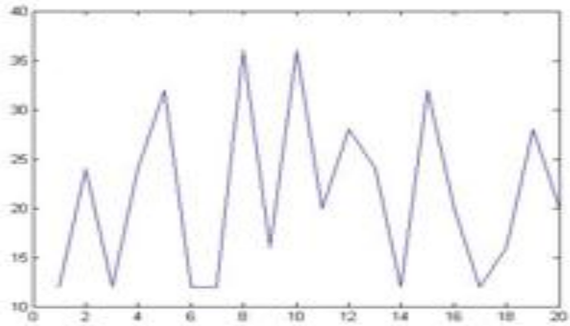

Fig 3. The power of each window in classification

Feature Extraction Using PCA

The Principle Component Analysis (PCA) is a classical well-structured method for feature extraction, which has been used several times in BCI researches [8]-[10].

PCA is generally used for dimensional reduction of the original data into its first d Eigen vectors (Fig.4) [11]-[14]. It converts a set of features of possibly correlated variables into a set of values of uncorrelated (orthogonal) variables called principal components which are minimal but sufficient. PCA causes a consistent increase in the speed of the classification, so it is a fast and useful method for extracting features [15],[16].

Feature Selection

The Extracted features are also relevant for classification and irrelevant which play no important role in the classification. An appropriate selection of features can actually improve the classifying and generalizing ability of the classifier, and can increase the speed and performance of the system. For selecting such features, there are different feature selection methods. Therefore, we use genetic algorithm and Particle Swarm Optimization separately to select the best combination of features.

The PCA algorithm

The initial data matrix should be supplied. first for each window of signal that is a matrix with several rows, convert it into a vector of M × 1 dimension by placing its rows side by side. Then place each vector in each column of an initial matrix X with M*N dimension. For applying PCA method on the initial matrix X, these steps were followed:

First, calculating the empirical mean (Eq.1):

$$U[m] = \frac{1}{N}\sum_{n=1}^{N} X[m,n] \qquad (1)$$

Second, calculating the standard deviations from the mean (Eq.2):

$$B = X - uh \text{ , h[n]=1 for n=1,…,N} \quad (2)$$

Third, finding the covariance matrix (Eq.3):

$$C = E[B \otimes B] = E[B.B^*] = \frac{1}{N}\sum B.B^* \quad (3)$$

Finally, calculating the eigenvectors and eigenvalues of the covariance matrix (Eq.4):

$$V^{-1} C V = D \quad (4)$$

The first principal component has as high a variance as possible, and each succeeding component in turn has the possibly highest variance. After rearranging the eigenvectors in order of decreasing eigenvalues, a subset of the top eigenvectors, as basis vectors, is selected and then the source data is converted to z-scores. Eq.5 and Eq.6 shows the creation of an M × 1 empirical standard deviation vector from the square root of each element along the main diagonal of the covariance matrix C:

$$S = \{s[m]\} = \sqrt{C[p,q]} \text{ for p=q=m=1,…,M} \quad (5)$$

$$Z = \frac{B}{s.h} \quad (6)$$

With projecting the z-scores of the data onto the new basis, the final transformed data is obtained (Eq.7). In Eq.7 W^* is the conjugate transpose of the eigenvector matrix [16].

$$Y = W^*.Z \quad (7)$$

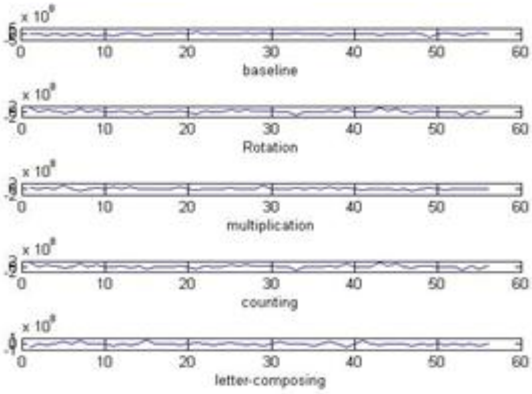

Fig 4. Feature Extraction Using PCA for EEG Signal of Five Mental Tasks

Genetic Algorithm (GA)

Genetic algorithm is an iterative method that examines different solutions in a large space to find the best solution. This method consists of four steps:

1. Producing initial population 2.Evaluation 3.Reproduction (selection and crossover) 4. Mutation.

In the first step, a population of chromosomes is generated. Each chromosome represents an independent solution. In this paper, each chromosome selects a singular combination of windows and new form of the signals made.

In the evaluation step, the power of each chromosome in the problem solving is to be examined. For this purpose, after forming new signals based on each chromosome, we classify test emails by KNN. The success rate of the classifier is saved as fitness value of that chromosome.

After evaluating chromosomes, in the selection step, half of the best chromosomes are selected based on their fitness values. With combining the best of chromosomes in crossover step, a new population will be created that is smarter than the previous one. The new population is replaced with the previous population and will be used in the next iteration of the algorithm. Commonly, the process terminated when either a maximum number of iterations achieved, or a satisfactory fitness value is obtained for one of the chromosomes. At the end of each period, the mutation step would be applied on the chromosomes. In this operation, a chromosome is randomly selected from the last population and one or more of its genes will be changed randomly. Therefore, the modified value of that chromosome usually does not exist in any of its parents. At the end of the optimization step, the best chromosome from the last population is selected and the new signals are formed based on that. The volume of the new signals is lower than the previous, but the power of the classifier has increased and the classifier is more efficient in time and performance [17].

Particle Swarm Optimization (PSO) for the EEG signal detection

The Particle Swarm Optimization (PSO) is a population based stochastic optimization technique that models behavior of birds within a flock. This method was applied successfully on a wide range of optimization problems [18], [4]. Similar to GA, PSO's aim is finding the optimal solution by searching complex space via collaboration and competition of particles and individuals. Particles within the swarm move in search space and learn from each other through local and global interactions [19]. When an appropriate solution found by one of the particles, all particles follow it in order to be more similar to the best particle. Different neighborhood types form different structures of the PSO, such as star topology, ring topology and wheels topology [20], [21]. For applying PSO, one should consider different combinations of windows as different solutions and construct initial particles. After making the final feature vector by PCA, we classify signals by KNN. The true positive of classifier are saved as fitness value of that particle.

In the original formulation of this algorithm [22], each particle defines a potential solution to the problem in a D-dimensional space. The particle i is represented in a D-dimensional space as:

$X_i = (x_{i1}, x_{i2}, x_{i3}, \ldots, x_{iD})$

In addition, the velocity for the i^{th} particle represented as:

$V_i = (v_{i1}, v_{i2}, v_{i3}, \ldots, v_{iD})$

The new position of each particle is introduced as follow:

$$V_k^i = \eta_1 * |R_1| * (X_{pBest} - X_k^i) + \eta_2 * |R_2| * (X_{gBest} - X_k^i) \quad (8)$$

$$X_{k+1}^i = X_k^i + V_k^i \quad (9)$$

Where η_1 and η_2 are positive acceleration constants and $|R_1|$ and $|R_2|$ are positive random numbers generated according to the absolute value of the Gaussian probability distribution, i.e. abs (N(0, 1)). V_i^k is the velocity vector of the i^{th} particle which simulates the optimization process and reflects the socially exchanged information [13],[23].

Each particle maintains a memory of its previous best position. X_{pbest} expresses the best previous position of the i^{th} particle, which has better fitness value ever. X_{gbest} represents the particle which has the best fitness value in all particles X^i_k. The optimized sampling process stopped if the best fitness value X_{gbest} reaches a certain threshold or the number of iterations gets to the certain value. After optimization, one can get the particles that are distributed around the best area [20], [21].

KNN classifier with different similarity measure

The K-Nearest Neighbor (KNN) is relatively a simple and discriminative nonlinear classifier. It assigns a class to a test feature vector according to its k nearest neighbors in training set [24]. The KNN is not very popular for EEG applications, maybe because they are known to be very sensitive to the high dimensional feature spaces. However, in a low dimensional feature space, similar to the feature space of this paper, maybe it is an efficient classifier [25]. In order to measure the distance in the KNN method, several similarity measures exist that define different types of the KNN. In this paper, four similarity measures are introduced and implemented and the power of them are compared. These distances are represented by Equations 10-13. In these formulas, q_i is related to the i^{th} component of the k^{th} feature vector of test data, and p_i represents i^{th} component of k^{th} feature vector of training data. For example, the metric Euclid Distance k_j, calculates the Euclid distance between k^{th} test data to the j^{th} training data. Another metrics are City Block, Chebishev and cosine metrics.

$$EuclidDistance_{kj} = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \tag{10}$$

$$Cityblock_{kj} = \sum_{i=1}^{n} | p_i - q_i | \tag{11}$$

$$Chebyshev_{kj} = MAX \ | p_i - q_i | \tag{12}$$

$$Cosine\_Metric_{kj} = Cos\theta = \frac{\sum_{i=1}^{n} p_i * q_j}{\sqrt{\sum_{i=1}^{n}(pi)^2} * \sqrt{\sum_{i=1}^{n}(qi)^2}} \tag{13}$$

Experimental Results

The proposed method was examined on the 325 EEG signals the 260 of which were used for training system and 65 of which for testing system. The power of the KNN, for different similarity measures were examined as they are shown in Fig.5. These results show that City block metric has the best results in classifying EEG signals by the KNN. So this metric is used in next parts of the proposed method.

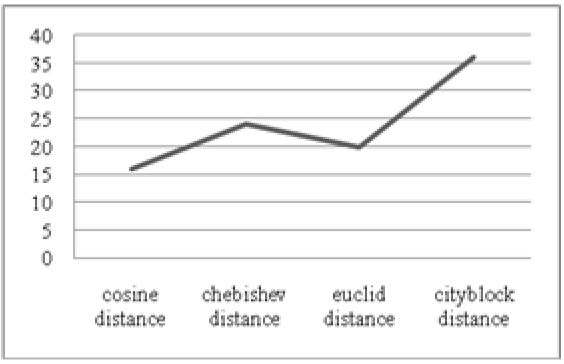

Fig 5. Comparing different similarity measures in KNN

In order to find the best combination of windows using GA as feature selection method, the true positive was increased by 8 percent. This method was examined with 12 chromosomes and with different number of iterations (epoch) 10, 20, 30, and 40. Fig.6 shows the chromosomes' average and maximum results in different periods. It is obvious from Fig.6 that the best results belong to the 20 and 40 periods.

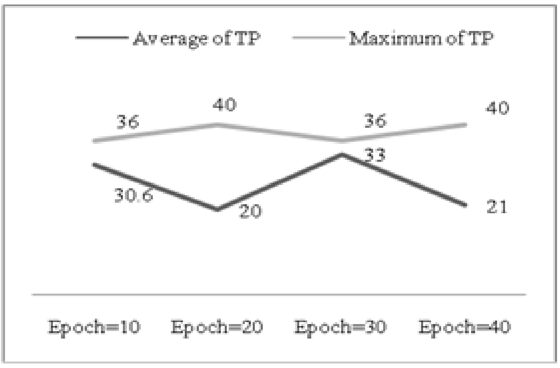

Fig 6. Comparing the TP of GA in different number of iterations

Another feature selection method used for finding the best combination of windows was PSO. This method, examined 12 particles in the 30 iterations. In Fig.7, the power of GA is compared with PSO. Fig.7 shows that GA algorithm always has got better results compared with PSO. Only in particle2, PSO shows the better results compared with GA.

Conclusion

In this research, EEG signal classification was examined. First, in preprocessing step we divided signals into several windows because the different parts of signal have different power in classification. The high volume of time components of EEG signals for PCA method is also another reason for partitioning signals. After partitioning the signals, PCA algorithm was used in order to extract some features from each window. For finding the best combination of windows in signals, two feature selection methods, GA and PSO, were used. The experimental results show that GA is a better method compared with PSO. GA increased the true positive of the classifier by 8 percent, but PSO increased it only by 4 percent. So, the best chromosome was selected with GA method, and the new signals were built based on that. These new signals are appropriate for classification.

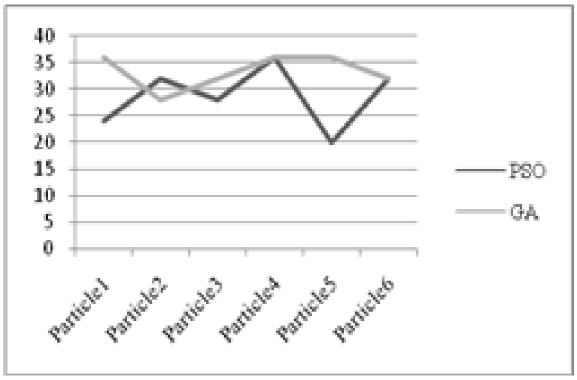

Fig 7. Comparing maximum success rate between PSO and GA

Acknowledgment

The authors would like to thank Dr Ali Ghahreman for editing the manuscript.

References

[1] E. Acar ,"classification of Motor Imagery Tasks In EEG Signal and ITS Application To a Brain-Computer Interface For Controlling Assistive Environmental Devices", Thesis for the degree of Master of Science, 2011.

[2] E. Niedermeyer, F.H.L.D. Silva, "Electroencephalography: basic principles, clinical applications, and related fields", Lippincott Williams & Wilkins, 2005.

[3] "Electroencephalography," http://en.wikipedia.org/wiki/Eeg, last visited on 15/01/2011.

[4] C.R. Hema, M.P. Paulraj, R. Nagarajan, S. Yaaco, A.H. Adom, "Application of Particle Swarm Optimization for EEG Signal Classification", Biomedical Soft Computing and Human Sciences, Vol.13, No.1, pp.79-84, 2008.

[5] M. Teplan, "Fundamentals of EEG Measurement", Measurement Science Review, Vol. 2, Section 2, 2002.

[6] J. Culpepper, "Discriminating Mental States Using EEG Represented by Power Spectral Density", Department of Computer Science Harvey Mudd College.

[7] Ch. W. Anderson, "correlating Brainwave Patterns with Mental Tasks Using MATLAB to Train Neural Networks on a Parallel Computer", a short paper in html, from CEBL laboratory of Colorado University.

[8] C. Lin, M. Hsieh, “Classification of mental task from EEG data using neural networks based on particle swarm optimization”, Neurocomputing, vol. 72, pp. 1121-1130, Jan. 2009.
[9] L. Ke, R. Li, “Classification of EEG Signals by Multi-Scale Filtering and PCA”, IEEE International Conference on Intelligent Computing and Intelligent Systems, pp. 362-366, 2009.
[10] H. Lee, S. Choi, “PCA-based linear dynamical systems for multichannel EEG classification” Proceedings of the 9th International Conference on Neural Information Processing, vol. 2, pp. 745-749.
[11] A. Subasi, I. Gursoy, “EEG signal classification using PCA, ICA, LDA and support vector machines”, Expert Systems with Applications 37 8659–8666, 2010.
[12] L.J. Cao, K.S. Chua, W.K. Chong, H.P. Lee, Q.M. Gu,” A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine”, Neurocomputing, 55, 321–336, 2003.
[13] C.R. Hema, M.P Paulraj, S. Yaacob, A.H. Adom, R. Nagarajan, “Motor Imagery Signal Classification for a Four State Brain Machine Interface”, World Academy of Science, Engineering and Technology 29, 2007.
[14] C.W. Anderson , S.V. Devulapalli ,A. Stolz, ”EEG Signal Classification with Different Signal epresentation”, IEEE , pp 475- 482, 1995.
[15] A. Lad, “Spam Net-Spam Detection Using PCA and Neural Networks”, CIT'04 Proceedings of the 7th international conference on Intelligent Information Technology, 2004.
[16]“Principle component analysis”, http:// www.wikipedia.com, last visited on 15/05/2012.
[17] G. Harik, F. Lobo, M. Kaufmann, “A Parameter-Less Genetic Algorithm”, Proceedings of the Genetic and Evolutionary Computation Conference, 258-265, 1999.
[18] I. Kenny, "DYNAMIC,HIERARCHICAL PARTICLE SWARM OPTIMISATION", Department of Computing Faculty of Mathematics and Computing The Open University Walton Hall, Milton Keynes United Kingdom, 2008.
[19] A. Atyabi, M. Luerssen, S.P. Fitzgibbon, D. Powers, “Adapting Subject-Independent Task-Specific EEGFeature Masks using PSOAdapting Subject-Independent Task-Specific EEG Feature Masks using PSO”, WCCI IEEE World Congress on Computational IntelligenceJune, Australi, 2012.
[20] ZH. You-ping, K. Xing-hong, H. Pei-wei, “Improved Algorithm for Distributed Localization in Wireless Sensor Networks”, Shanghai Jiaotong University and Springer-Verlag Berlin Heidelberg, 2010.
[21] A.P. Engelbrecht, J. Wiley, “Computational Intelligence An Introduction”, 2002.
[22] V. Bostanov, “BCI Competition 2003- Data Set Ib and IIb: Feature Extraction from event Related Brain Potentials with the Continuos Wavelet Transform and t-Value Scalogram”, IEEE Trans. Biomedical Engineering, pp 1057- 1061, 2004.
[23] Y. Yuhui Shi, R.C. Eberhart, “Parameter Selection in Particle Swarm Optimization”, Evolutionary Programming, Springer Verlag, New York , Vol. 7 , pp 591 – 600, 1998.
[24] S. Bhattacharyya, A. Khasnobish, S. Chatterjee, A. Konar, D.N. Tibarewala, “Performance Analysis of LDA, QDA and KNN Algorithms in Left-Right Limb Movement Classification from EEG Data”, International Conference on Systems in Medicine and Biology, liT Kharagpur, India, 2010.
[25] F. Lotte, M. Congedo, A. ecuyer, F. Lamarche, B. Arnaldi, “A Review of Classi_cation Algorithms for EEG-based Brain-Computer Interfaces”, inria-00134950, version 1– 6, 2007.