Multi-Objective Binary-PSO for Improving the Performance of P300 Speller

Rahul Kumar Chaurasiya[2*], Aparajita Saraf[1],D. Sarala [2], TejasKanikdaley[2], SharadJogi [2],

[1]Dept. of ECE, University of Michigan, USA
[2]Dept. of Electronics and Telecommunication,National Institute of Technology Raipur, India
[*]Corresponding-author-rkchaurasiya@nitrr.ac.in

***Abstract*—P300 speller provides a medium of communication to patients with severe neuro-muscular disabilities. It relies on the detection of P300 peak in EEG recordings due to the occurrence of external stimulus. The patients are asked to focus their attention on the character/symbol displayed on the screen and their EEG recordings are monitored at the same time. These recorded signals are then processed and analyzed to determine the target character. Thus the patient is able to spell out the word he intended to communicate. In the process, a large amount of redundant and noisy data is also collected which reduces the overall accuracy of the speller.**

This paper proposes the use of multi-objective binary particle swarm optimization (MO-BPSO) technique to identify the most relevant electrode channels on an EEG headset for data collection. The trade-off between the accuracy of the character detection and the number of channels is determined using the pareto-optimal solution plot. The user can later select any of the pareto-optimal solution based on his requirement. The dataset used for this purpose is of BCI competition III. The result section also discusses the channel and thepositioning which play a major role in P300 signal detection.

***Keywords*—P300 speller, optimal channel selection, EEG, PSO, multi-objective optimization, SVM.**

I. INTRODUCTION

Brain-computer interfaces (BCI) are communication pathways which allow human beings to interact with the computer without involving any muscular activity. The brain directly communicates with the external devices.BCI devices have proved effective for locked-in and paralyzed patients who have lost control over their limbs[1].BCI incorporates researchers from the fields of neuroscience, engineering, physiology, psychology, medical health care and robotics. The applications of BCI spans to rehabilitation, game playing, brain controlled drone racing, emotion detection, sleep detection, brain fingerprinting, and concentration improvement exercises[2-5].

There are two methods of recording brain signals: invasive and non-invasive. In invasive BCI system, the electrodes are placed inside the brain to record the signals whereas in non-invasive method, electroencephalography (EEG) headset is used for recording signals. Though invasive systems provide better signal to noise ratio, non-invasive BCI systems are preferred for their less risky and convenient signal acquisition techniques[6].

P300 speller, works on non-invasive BCI system, which was first introduced in 1988 by Farewell and Donchin[7]. P300 speller is based on the oddball paradigm, where an infrequent stimulus produces a positive deflection in EEG recordings after 300ms[8].P300 speller row/column (RC) paradigm consists of a matrix of English alphabets and numbers. The rows and columns of this matrix are randomly intensified. The subject is asked to focus on the character displayed on the matrix that he wants to spell (target character). When the row/column containing the target character is intensified, a peak is observed in the EEG recordings of the subject. This peak in EEG signals is used to detect target character. Since signal to noise ratio is low, character detection from single trial is difficult[8].For improving accuracy in character detection, multiple trials are done for the same character.

There are many problems associated with the P300 speller which leads to incorrect deduction of the character in the matrix. Adjacency[9], crowding effect[10, 11], habituation[12] and fatigue[13] deteriorate the performance of the speller. Also, there is considerable amount of irrelevant and redundant data present in the EEG recordings. If we can find the relevant channels which together give a better result, we can get rid of irrelevant data, reduce the computation time anduse configuration of these channels for future recordings. To maximize the accuracy in character detection and minimize the number of channels, we propose to employ multi-objective particle swarm optimization (MOBPSO)-based approach. There are various evolutionary optimization techniques that have been studied but PSO was chosen for its simpler implementation, faster computational speed and an efficient global optimizer quality[14].The solutions are chosen for the pareto-front such that one objective of any data point cannot be increased without compromising at least one another objective. Thus, all the solutions in the pareto-front are considered optimal and different solutions are selected according to the requirement.

The rest of the paper structure is as follows: Section 2 briefly describes the related work in P300 speller. Section 3 elaborates the data set, pre-processing and classification methods used. Section 4 describes the proposed optimization techniques. Section 5 compares the result of the MOBPSO technique against existing methods; followed by conclusion in Section 6.

II. RELATED WORK

Several variants have been proposed for BCI-based P300 speller. P300 speller with different display paradigms[15-17], with characters in different languages[18-20], with images instead of characters[21-23], with different colours and font size[9, 24] and with different sequence of random intensifications[25] have been proposed. Several classification

978-1-5090-6785-5/18/$31.00 © 2018 by IEEE

methods such as support vector machines (SVM), artificial neural network (ANN), Bayesian linear discriminant analysis (BLDA) and step-wise linear discriminant analysis (SWLDA) have successfully classified target vs. non target stimuli [26-32]. Various optimization techniques have also been used for better channel selection like particle swarm optimization (PSO), genetic algorithm (GA) and differential evolution (DE) [32-34].

III. MATERIALS AND METHODS

A. Dataset

The data set is Wadsworth BCI dataset for P300 evoked potentials from BCI Competition III (2004)[35]acquired using a general purpose software named BCI2000[36]with P300 Speller paradigm shown in Figure 1. The signals were collected from two subjects in a session of five each, which was then filtered through a band-pass filter (0.1- 60Hz) and then digitized at 240Hz. The given datasethas train and testmatrix for both the subjects.

Fig. 1.The 6×6 P300 display paradigm used for recording the dataset.One of the rows is intensified in the presented image.

B. Pre-Processing

Pre-processing is done to remove noise from the data and to improve the signal-to-noise ratio (SNR). Data sample are collected from 0 to 600ms after the each intensification. As P300 ERP is generated after around 300ms, the time period of 600ms is sufficient to analyse the signals. Since, only P300 peak has to be detected for classification purposes, we need only low frequency component of the data which contains the basic information of EEG signals. These extracted data points are then passed through a band-pass fourth order Chebyshev Type I filter having a cut-off frequency of 1 to 10Hz. The filtered signals are decimated to cut-off frequency of 10Hz.Then normalization is carried out by subtracting each data point by its mean and dividing it by the standard deviation of the corresponding channel. Figure 2 visualizes two different signals, one with +1 class label (i.e. with P300 ERP signals) and another with -1 class label (i.e. without P300 ERP signals).

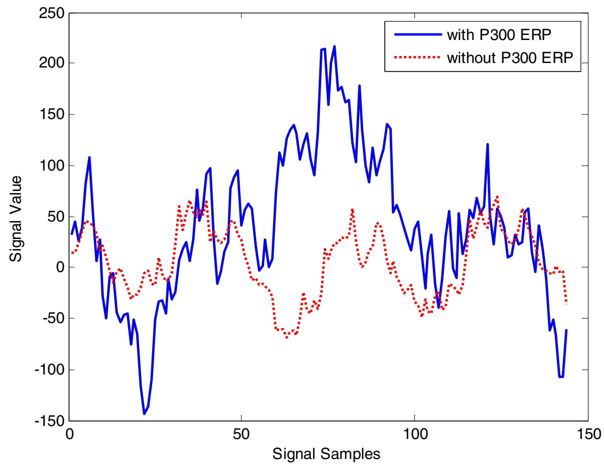

Fig.2.Raw signals with and without P300 ERP.

After pre-processing, the signal sampleshave been converted into vectors of 6 samples per channelper intensification.Thus, the training data is a 15300×384 matrix; 12 intensifications × 15 trials × 85 characters in training data set =15300 and 6 samples × 64 channels = 384 samples. The pre-processing is repeated for the test data where the final data set after pre-processing is of size18000×384 (there are 100 characters in test data set: 12 intensifications × 15 trials × 100 characters in training data set =18000).

C. Classification using SVM-based Classifier

The classification process is divided into training and testing stages. In the first stage, the SVM system is trained with the trainingdata provided in the dataset and then in the second stage, the SVM system maps the test data to the class labels. The accuracy is calculated by comparing the actual labels against the labels allotted by the SVM system.SVM-based classification is chosen because it has better generalization capabilities over other classification methods[37].

The SVM system is trained by using *averaged feature vectors* of the training matrix. There are 12 intensifications per character and for its correct detection; the user focuses his attention on the same target character for 15 trials (i.e. for 12×15 =180 intensifications). The P300 peak signals of EEG recordings of same character are averaged. This reduces the effect of noise and the data to be processed is significantly compacted.The SVM system is trained on this reduced data set. It learns a hyper plane $\omega^T x + b = 0$that maximizes the separation margin between the data points of two classes[38]. The learned hyper-plane is then used to assign a class label -1 or +1 to the test data points as in Eq. (1). Here, +1 denotes that the intensification of the corresponding row/column contains the target character.

D. Character Detection

After the SVM has allotted labels to the test data, for every one character, there are ten class -1 vectors and two class+1vectors (corresponding to one row and one column). The information of the row and column is used to detect the target character. The score vector S containing 12 score values, refers to the score values of 12 rows and columns for one character is given as:

$$S(j) = \sum_{i=1}^{N} y_i \lambda_i \boldsymbol{X_i} \boldsymbol{X_j} \qquad (1)$$

where, λ_ii =1,2…N are the Lagrange's multipliers. λ_i's zero for all points which are not support vectors.$\boldsymbol{X_i}$represents the training points for i= 1,2…N and $\boldsymbol{X_j}$ represents the test points corresponding to 12 different rows and columns for j=1,2…12. The correct column and row are predicted on the basis of the score values S as computed in Eq. (2) and Eq. (3).

$$\text{Predicted column} = \underset{1 \le j \le 6}{\arg max} S(j) \qquad (2)$$

$$\text{Predicted row} = \underset{7 \le j \le 12}{\arg max} S(j) \qquad (3)$$

Once the row and column are identified, the target character is identified from the display paradigm.

IV. Proposed MOBPSO-based Channel Selection

A. Binary PSO

Particle swarm optimization (PSO) was introduced in 1995 by Kennedy and Eberhart[39]. It is an evolutionary optimization technique inspired from the social behavior of bees, birds and fishes. PSO technique attempts to maximize a fitness function by moving a swarm of particles around the space defined by the fitness function and its constraints.

Initially, the particles are spread randomly in the space. The position of various particles is updated iteratively by its velocity. Each particle maintains a record of its previous best performance position asp_{best}.There are three basic parameters responsible for updating the velocity of a particle: personal best position p_{best}, global best position g_{best} and inertia. The velocity v_i of particle x_i is updated as:

$$v_i^{(t+1)} = w v_i^{(t)} + c_1 \phi_1 (p_{best} - x_i^{(t)}) + c_2 \phi_2 (g_{best} - x_i^{(t)}) \quad (4)$$

Further, the position of particle x_i for next iteration is computed as:

$$x_i^{(t+1)} = v_i^{(t+1)} + x_i^{(t)} \qquad (5)$$

where,w is the weight coefficient, $c_1$and c_2 are constants and generally $c_1 + c_2 = 4$ [40].The numbers ϕ_1 and ϕ_2 are generated randomly from a uniform distribution over 0 to 1.

Different coefficients $(w, c_1, c_2, \phi_1, \phi_2)$decide how much a particle would be affected by its own experience, experience of other particles and its inertia.This conventional version of PSO cannot be used for the problems defined in discrete domain. Hence, a binary particle swarm optimization (BPSO) was proposed.Instead of using real values for defining the position and velocity of the particle, BPSO uses only 0s and 1s[41].In order to integrate this change, a modification was done in the velocity term. In BPSO, the velocity term (v_i)refers to the probability that the particle (x_i)will change its position state to 1. Thus, the velocity (v_i) is restricted between 0 and1as:

$$v_i' = sig(v_i) = \frac{1}{1 + e^{-v_i}} \qquad (6)$$

Following this, the position of the particle is updated as:

$$x_i^{(t+1)} = \begin{cases} 1, & if\ rand() < sig(v_i) \\ 0, & otherwise \end{cases} \qquad (7)$$

This first version of BPSO may suffer from the problem of trapping into local minima. Hence, its variants are proposed. We have used the algorithm of [42]in this article.

B. MOBSO for Optimizing the Trade-Off between The Classification Accuracy and Number of Channels

BPSO and PSO were designed to workfor only one objective function. However, maximization of classification accuracy and minimization of number of channels is a multi-objective optimization problem. In this article, the positions of particles define the number of channels selected. The MOBPSO proposed in the article handles the multi-objective problem, which is mathematically formulated as follows:

$$\text{Objective 1: } \max(A(x_i))$$

$$\text{Objective 2: } \min(C(x_i))$$

$$s.t.\ x_i \subseteq x \qquad (8)$$

Here,A (x_i) and $C(x_i)$ are the accuracy of character detection and the total number of channels corresponding to particlex_i, respectively. The particlex_i is represented as a 1×64 dimensional matrix and it uses only those channels for classification, which are taken as 1 in the particle x_i.Mathematically, the channel$j$ is chosen, if $x_{ij} = 1$ for $j = 1$ to 64. The MOBPSO algorithm plots optimal solutions in the pareto-front, anyoptimal solution may be chosen depending on the application.

V. Results and Discussions

A. Results Obtained using SVM Classifier for The Data Collected from All 64 Channel

The classification accuracies obtained by SVM-based classifier for different number of trials are depicted in Table.1. For obtaining these results,data from all 64 channels were used for analysis.

TABLE I. Classification accuracies for different number of trials.

Subject	No. of trials		
	5	10	15
Subject A	70.8	82.1	91.0
Subject B	65.4	78.1	85.9

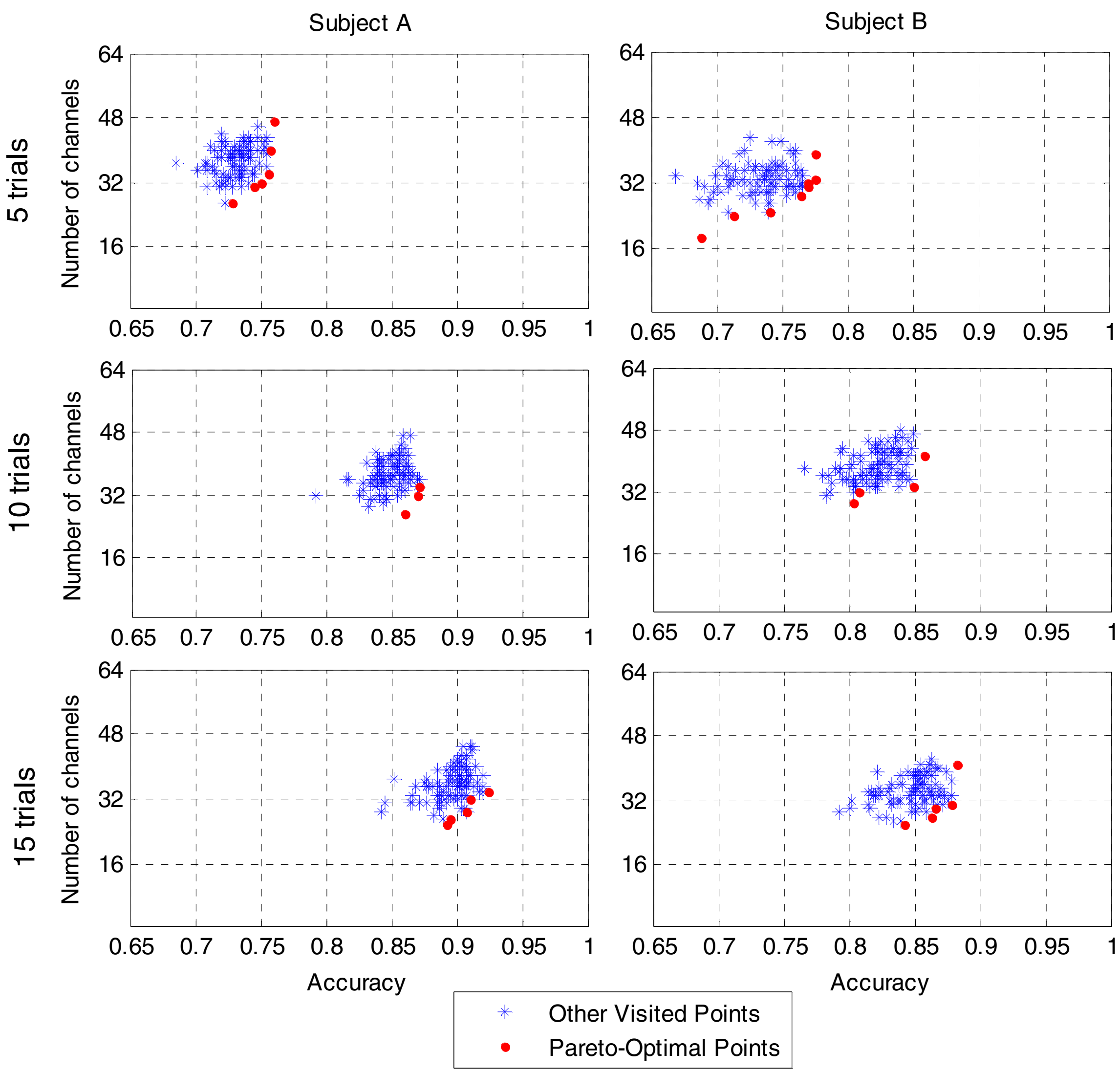

Figure 3: Pareto-optimal solutions for different trials

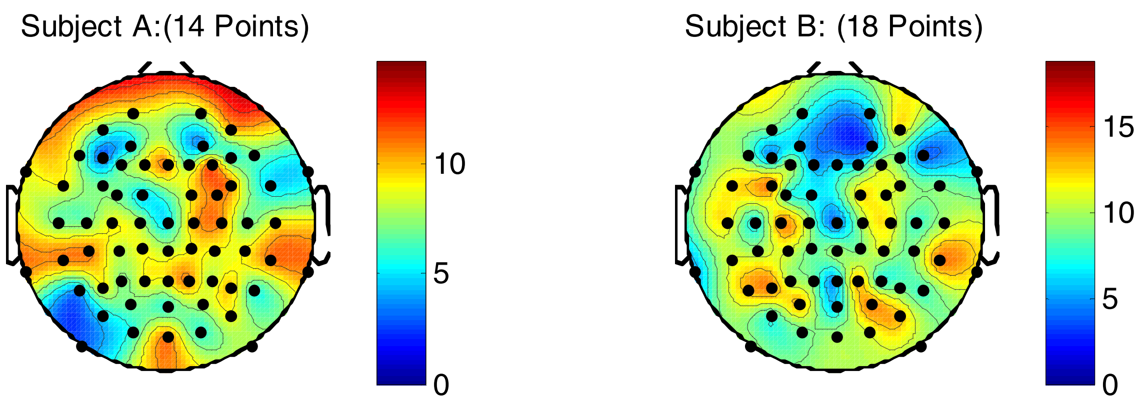

Figure 4: Channel topo-plot for two subjects

B. Results Obtained using SVM Classifier with BPSO-based Method for Maximizing The Classification Accuracy

The classification accuracies obtained for different number of trials after applying BPSO for maximizing the classification accuracy are tabulated in Table 2.The accuracy values have improved and the number of channels, required for processing, has decreased.

TABLE II. ACCURACIES FOR DIFFERENT NUMBER OF TRIALS FOR CLASSIFICATION WITH BPSO

Subjects	**Averaged accuracy (A) and number of channels selected(C) with different number of trials**					
	5 trials		**10 trials**		**15 trials**	
	A	**C**	**A**	**C**	**A**	**C**
Subject A	72.3	32	83.5	32	93.1	36
Subject B	70.5	29	79.1	31	88.4	25

C. Pareto-Pptimal Solutions using MOBPSO Algorithm

The aim of MOBPSO in this article is to attain two objectives are namely, to improve the classification accuracy and to decrease the number of channels. The pareto-front diagrams for different number of trials are shown in Figure 3.

D. Channel Topology

The various channel numbers and their placement is represented inFigure 5.

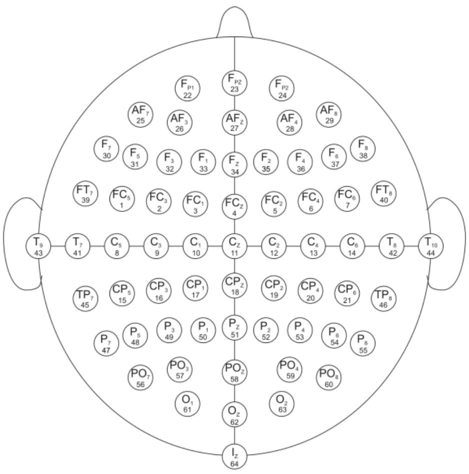

Figure 5: Channel configuration

Channel topoplot effectively displays the best set of channels for both the subjects. The data points obtained in pareto-front are used to determine most frequently used channels. The channel topoplot for the optimal solutions obtained by SVM-MOBPSO technique is visualized in Fig.4.

E. Statistical Analysis

Friedman test is often used to provide statistical comparison between various algorithms. In this test, the Friedman test has ranked different approaches according to their performance on different datasets[43]. The best approach for a given dataset is given 1st rank. The second best approach is selected for the 2nd rank, and so on. This procedure is repeated for different datasets. Then, the Friedman test statistics are calculated on the average rank obtained by the approaches. The Friedman test in this article was applied to two different classification approaches namely, SVM (trained by considering data from all 64 channels) and SVM used with BPSO optimization technique. The calculatedp-value (0.014306) was less than 0.05. The null hypothesis was rejected, following which, the post-hoc Nemenyi test was applied to report any significant difference between the individual approaches [44]. The results of the test are shown in Figure 6.

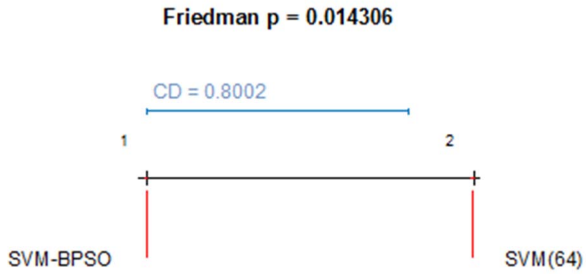

Figure 6: Results of Freidman test depicting SVM-BPSO's better performance over SVM(64).

As the rank difference between the two methods is greater than the critical distance, the first ranked method is significantly better than the second ranked method. Thus, theproposed methodwill always perform better than SVM.

VI. CONCLUSION

In this paper, MOBPSO technique was employed with SVM for optimizing the trade-off between classification accuracy and the number of channels used to collect EEG data. The proposed method has been tested and validated on the publicly available dataset. The results have provided pareto-optimal solutions; the user can later choose the solution according to his needs and requirements. Statistical analysis shown in the results, prove that the SVM with MOBPSO provides significantly better results than SVM alone.

In thepaper, the filtered and decimated amplitudes of EEG signals were considered as feature vectors. The future work is proposed to explore more features to further improve the performance of the system. Also, the work portrayed here was analyzed offline and its performance on online systems would be examined in future.

References

[1]J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin Neurophysiol,* vol. 113, pp. 767-91, Jun 2002.

[2]T. M. Vaughan, W. J. Heetderks, L. J. Trejo, W. Z. Rymer, M. Weinrich, M. M. Moore*, et al.*, "Brain-computer interface technology: a review of the Second International Meeting,"

978-1-5090-6785-5/18/$31.00 © 2018 by IEEE

IEEE Trans Neural Syst Rehabil Eng, vol. 11, pp. 94-109, Jun 2003.

[3]A. Finke, A. Lenhardt, and H. Ritter, "The MindGame: a P300-based brain-computer interface game," *Neural Netw,* vol. 22, pp. 1329-33, Nov 2009.

[4]V. Bostanov, P. M. Keune, B. Kotchoubey, and M. Hautzinger, "Event-related brain potentials reflect increased concentration ability after mindfulness-based cognitive therapy for depression: a randomized clinical trial," *Psychiatry Res,* vol. 199, pp. 174-80, Oct 30 2012.

[5]L. A. Farwell, "Brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials," *Cogn Neurodyn,* vol. 6, pp. 115-54, Apr 2012.

[6]E. A. Curran and M. J. Stokes, "Learning to control brain activity: a review of the production and control of EEG components for driving brain-computer interface (BCI) systems," *Brain Cogn,* vol. 51, pp. 326-36, Apr 2003.

[7]L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr Clin Neurophysiol,* vol. 70, pp. 510-23, Dec 1988.

[8] J. Elshout, "Review of brain-computer interfaces based on the P300 evoked potential," 2009.

[9] M. Salvaris and F. Sepulveda, "Visual modifications on the P300 speller BCI paradigm," *Journal of neural engineering,* vol. 6, p. 046011, 2009.

[10]H. Strasburger, "Unfocussed spatial attention underlies the crowding effect in indirect form vision," *Journal of Vision,* vol. 5, p. 8, 2005.

[11]C. Feng, Y. Jiang, and S. He, "Horizontal and vertical asymmetry in visual spatial crowding effects," *J Vis,* vol. 7, pp. 13 1-10, 2007.

[12]J. Jin, B. Z. Allison, E. W. Sellers, C. Brunner, P. Horki, X. Wang, *et al.*, "Optimized stimulus presentation patterns for an event-related potential EEG-based brain–computer interface," *Medical & biological engineering & computing,* vol. 49, pp. 181-191, 2011.

[13]J. Polich, "Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli," *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section,* vol. 68, pp. 311-320, 1987.

[14]J. Kennedy, "Particle swarm optimization," in *Encyclopedia of Machine Learning*, ed: Springer, 2010, pp. 760-766.

[15]C. Guan, M. Thulasidas, and J. Wu, "High performance P300 speller for brain-computer interface," in *Biomedical Circuits and Systems, 2004 IEEE International Workshop on*, 2004, pp. S3/5/INV-S3/13-16.

[16]G. Townsend, B. LaPallo, C. Boulay, D. Krusienski, G. Frye, C. Hauser, *et al.*, "A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns," *Clinical Neurophysiology,* vol. 121, pp. 1109-1120, 2010.

[17]R. Fazel-Rezai and K. Abhari, "A region-based P300 speller for brain-computer interface," *Electrical and Computer Engineering, Canadian Journal of,* vol. 34, pp. 81-85, 2009.

[18] J. Jin, B. Z. Allison, C. Brunner, B. Wang, X. Wang, J. Zhang, *et al.*, "P300 Chinese input system based on Bayesian LDA," *Biomed Tech (Berl),* vol. 55, pp. 5-18, Feb 2010.

[19]Y. Yamamoto, T. Yoshikawa, and T. Furuhashi, "Improvement Of Performance Of Japanese P300 Speller By Using Second Display," *Journal of Artificial Intelligence and Soft Computing Research,* vol. 5, pp. 221-226, 2015.

[20]J. W. Minett, H.-Y. Zheng, M. C. Fong, L. Zhou, G. Peng, and W. S. Wang, "A Chinese text input brain–computer interface based on the P300 speller," *International Journal of Human-Computer Interaction,* vol. 28, pp. 472-483, 2012.

[21]J. D. Bayliss and D. H. Ballard, "A virtual reality testbed for brain-computer interface research," *Rehabilitation Engineering, IEEE Transactions on,* vol. 8, pp. 188-190, 2000.

[22]H. Serby, E. Yom-Tov, and G. F. Inbar, "An improved P300-based brain-computer interface," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on,* vol. 13, pp. 89-98, 2005.

[23]T. Kaufmann, S. M. Schulz, A. Köblitz, G. Renner, C. Wessig, and A. Kübler, "Face stimuli effectively prevent brain–computer interface inefficiency in patients with neurodegenerative disease," *Clinical Neurophysiology,* vol. 124, pp. 893-900, 2013.

[24]B. Z. Allison and J. A. Pineda, "Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system," *International journal of psychophysiology,* vol. 59, pp. 127-140, 2006.

[25]S.-K. Yeom, S. Fazli, and S.-W. Lee, "P300 visual speller based on random set presentation," in *Brain-Computer Interface (BCI), 2014 International Winter Workshop on*, 2014, pp. 1-2.

[26]V. Bhatnagar, N. Yede, R. S. Keram, and R. Chaurasiya, "A modified approach to ensemble of SVM for P300 based brain computer interface," in *2016 International Conference on Advances in Human Machine Interaction (HMI)*, 2016, pp. 1-5.

[27]H. Bhagwani and R. K. Chaurasiya, "Step Wise LDA on Averaged ERPs for Faster and Efficient Brain-Computer Interfacing," in *2nd IEEE International Conference on Human Computer Interactions (ICHCI'16)*, Chennai, India, 2016, pp. 26-29.

[28]R. K. Chaurasiya, N. D. Londhe, and S. Ghosh, "An Efficient P300 Speller System for Brain-Computer Interface," in *IEEE International Conference on Signal Processing, Computing and Control (ISPCC)*, JUIT Solan India, 2015.

[29]A. Rakotomamonjy and V. Guigue, "BCI competition III: dataset II- ensemble of SVMs for BCI P300 speller," *IEEE Trans Biomed Eng,* vol. 55, pp. 1147-54, Mar 2008.

[30]R. Fazel-Rezai, B. Z. Allison, C. Guger, E. W. Sellers, S. C. Kleih, and A. Kubler, "P300 brain computer interface: current challenges and emerging trends," *Front Neuroeng,* vol. 5, p. 14, 2012.

[31]M. Akcakaya, B. Peters, M. Moghadamfalahi, A. R. Mooney, U. Orhan, B. Oken, *et al.*, "Noninvasive brain-computer interfaces for augmentative and alternative communication," *IEEE Rev Biomed Eng,* vol. 7, pp. 31-49, 2014.

[32]R. K. Chaurasiya, N. D. Londhe, and S. Ghosh, "Binary DE-Based Channel Selection and Weighted Ensemble of SVM Classification for Novel Brain–Computer Interface Using Devanagari Script-Based P300 Speller Paradigm," *International Journal of Human–Computer Interaction,* vol. 32, pp. 861-877, 2016/11/01 2016.

[33]J. Mak, Y. Arbel, J. Minett, L. McCane, B. Yuksel, D. Ryan, *et al.*, "Optimizing the P300-based brain–computer interface: current status, limitations and future directions," *Journal of neural engineering,* vol. 8, p. 025003, 2011.

[34]C. Y. Kee, R. K. Chetty, B. H. Khoo, and S. Ponnambalam, "Genetic Algorithm and Bayesian Linear Discriminant Analysis Based Channel Selection Method for P300 BCI," in *Trends in Intelligent Robotics, Automation, and Manufacturing*, ed: Springer, 2012, pp. 226-235.

978-1-5090-6785-5/18/$31.00 © 2018 by IEEE

[35]B. Blankertz, K. R. Muller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl*, et al.*, "The BCI competition. III: Validating alternative approaches to actual BCI problems," *IEEE Trans Neural Syst Rehabil Eng,* vol. 14, pp. 153-9, Jun 2006.

[36]G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *Biomedical Engineering, IEEE Transactions on,* vol. 51, pp. 1034-1043, 2004.

[37]A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 22, pp. 4-37, 2000.

[38]S. Theodoridis and K. Koutroumbas, *Pattern Recognition, Fourth Edition*: Academic Press, 2008.

[39]R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the sixth international symposium on micro machine and human science*, 1995, pp. 39-43.

[40]G. D. A. Carlisle, "An off-the-shelf PSO," presented at the Proceedings of the Workshop on Particle Swarm Optimization, January 2001.

[41]J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on*, 1997, pp. 4104-4108.

[42]S. Mirjalili and A. Lewis, "S-shaped versus V-shaped transfer functions for binary particle swarm optimization," *Swarm and Evolutionary Computation,* vol. 9, pp. 1-14, 2013.

[43]J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *The Journal of Machine Learning Research,* vol. 7, pp. 1-30, 2006.

[44]P. Nemenyi, "Distribution-free multiple comparisons," in *Biometrics*, 1962, pp. 263-&.

978-1-5090-6785-5/18/$31.00 © 2018 by IEEE