2014 Middle East Conference on Biomedical Engineering (MECBME)
February 17-20, 2014, Hilton Hotel, Doha, Qatar

Electrode Reduction Using ICA and PCA in P300 Visual Speller Brain-Computer Interface System

Abeer E. Selim, Manal Abdel Wahed, and Yasser M. Kadah

Abstract— **Brain-Computer Interface (BCI) research aims at developing systems helping disabled people hereafter called subjects. Due to the fact that technology underlying BCI is not yet mature enough and still having shortcomings for usage out of laboratory, these prevent their widespread application. These shortcomings are caused by limitations in functionality of BCI system tools and techniques. The motivation of this work was to develop efficient BCI techniques including signal processing, feature extraction, pattern recognition and classification to improve the performance of P300 Visual Speller BCI system. Data sets used in this paper were acquired using BCI2000's P300 Speller paradigm provided by BCI competitions. Primarily, in the processing phase time domain and spatial domain feature extraction were applied. Followed by classification phase where various linear and extended linear classifiers were utilized. One of the main achievements of this paper is applying Independent Component Analysis (ICA) or Principal Component Analysis (PCA) as spatial domain feature extraction for dimensionality and artifact reduction. Reducing electrodes to half its original size highly improved performance with linear classifiers and yet outperformed the results of BCI competition winners with extended linear classifiers.**

I. Introduction

The ability to communicate with others through speech, gesturing, or writing is one of the main factors facilitating human lives. These means of communication can help people expressing their ideas, desires, and feelings, allowing them to cope with daily life tasks. Some people suffer from sever motor disabilities due to some neurological diseases. They are fully conscious but unable to communicate with others or produce any motor output. There must be an alternative technology assisting them to communicate with their environment. The only alternative may be in exploring indirect voluntary modulation of electrical fields resulting from neural processes in their brain activity. BCI systems are promising mean that give back communication abilities. The idea underlying BCI is measuring electric, magnetic, or other physical manifestations of brain activity and detecting patterns of brain activities then translating these patterns into commands for a computer application or other devices. The basic model of BCI system consists of three main components. First component is brain activities representing adaptive controller, second is BCI tools and techniques considered as the core of BCI systems translating brain activities into commands or control signals and last component is the application or the device to be controlled.

In order to employ brain activities they have to be manipulated by subjects via special thoughts. Subjects have to acquire conscious control over their brain activities to produce discriminant patterns. BCI systems offer different paradigms to help subjects manipulating their brain signals and consequently, different brain activity patterns can be obtained. In specific cases paradigms have to be chosen depending on subject's abilities, willingness and application scenario. Examples of paradigms are systems presenting group of symbols to choose one, concentrating on a specific mental task such as mathematical operations, controlling movement of artificial limb by imaginary movements or Visual Speller helping subjects to spell words which is the paradigm of interest in this paper. Different paradigms can yield two types of brain activities. First, is stimulus-driven where visual or auditory stimulus is sent to subject and yield evoked brain activity. No subject training is needed, training load will be on BCI system to recognize brain evoked patterns. Second, is user-driven where subject should be trained to produce easily detectable spontaneous brain activity patterns. Feedback signals are often sent back to train subject, to improve their brain activity. Therefore, the training load is mainly on subject, not on BCI system.

The most popular, safe and widely used method of acquiring brain activities in BCI systems is the Electroencephalogram (EEG). EEG acquisition devices are relatively inexpensive, easily transported and fast setup. However, the disadvantages of EEG signals are their high dimensionality, multichannel nature, low signal-to-noise ratio, and being artifact contaminated. EEG artifacts could be physiological such as eye movement/blinks, muscles/heart activities or non-physiological including power supply, EEG amplifier and electrode–scalp interface noise.

Associated to BCI paradigms, the problem of classifying these patterns which is a challenging problem due to EEG signal nature described above. Moreover, the high variability between different subjects and the variance in subject performance. Most BCI systems contains as a core part supervised Machine Learning Algorithms (MLAs), learning from training datasets to discriminate different brain activity patterns in testing datasets. MLAs adapt BCI system to the brain of a particular subject which decreases learning load on

Abeer E. Selim, GDC (Global Delivery Center), IBM, Egypt (email:abeers@eg.ibm.com / abeerselim.as@gmail.com)
Manal Abdel Wahed, Biomedical Engineering Department, Cairo University, Egypt (email: manal.wahed@eng.cu.edu.eg)
Yasser M. Kadah, Biomedical Engineering Department, Cairo University, Egypt (email: ymk@k-space.org).

978-1-4799-4799-7/14/$31.00 ©2014 IEEE

subjects. MLA performance depends on feature extraction and classification techniques employed. The reliable results of MLAs can be used to establish group of commands controlling variety of devices or computer applications. Theoretically, any device can be connected to a computer or a microcontroller could be controlled with BCI. However, in practice devices and applications controlled with BCI are limited such as wheelchair and artificial limb.

I. BRAIN ACTIVITY

A. *Brain Signal*

Since few years, several BCI competitions were organized to promote the development of BCI and evaluate current state of the art of BCI system tools and techniques. Well versed laboratories in EEG-based BCI research provided datasets in a documented format. These datasets classified into labeled training sets and unlabeled testing sets. The competition goal was to maximize the performance measure for test labels. These competitions allow the community to benchmark several classification techniques in an unbiased way [8]. Datasets used in this paper acquired using BCI2000's P300 Speller paradigm [1] provided by BCI competitions II (2003) [2] recorded from subject ‘C’ and competition III (2004) recorded from subjects ‘A’ and ‘B’. These datasets represent a complete record of P300 Evoked Related Potentials (ERPs) [3].

B. *BCI Paradigm and Brain Activity Manipulation*

P300 Visual Speller based on the so-called oddball paradigm which states that rare expected stimuli produce change in brain activity. In this paradigm subject is presented by 6x6 matrix of characters *Fig. 1* [7]. The matrix is encoded from 1-12 *Fig. 2* [8]. The subject’s task is to spell the word displayed at top of the matrix, one character at a time.

During spelling task brain signals are acquired using 64 electrodes EEG equipment. For spelling of a single character, each of the 12 (6 rows/columns) in the matrix successively and randomly is intensified. Two particular row and column out of 12 contain the desired character *Fig. 1*. Detecting which stimulus the subject is concentrating upon is equivalent to detecting which stimulus caused manipulation of brain activities in the corresponding EEG segment. Brain manipulation termed Event Related Potential (ERP). ERP is the most widely used neurophysiologic activity to derive BCI systems. The most robust component of ERP is P300 which is a positive deflection appearing in brain signals within approximately 300ms after the presentation of a task-significant or noteworthy stimulus within random series of stimuli. P300 automatically and involuntarily appears almost in all people. This BCI system is suitable for subjects who might have difficulties in acquiring voluntary control on their brain activity, having concentration problems, or not willing to go through long training. In order to make spelling procedure more reliable, sets of 12 row/column intensifications repeated for each character 15 sequences. Recognition rate of spelled characters is the evaluation criteria. The goal is to correctly predict the desired character at fewest sequences as possible.

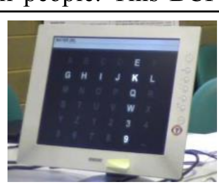

Figure 1. P300 Visual Speller [7].

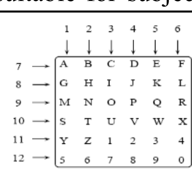

Figure 2. Encoding of matrix Rpws & Columns [8].

II. BCI TOOLS AND TECHNIQUES

BCI tools and techniques including signal processing, feature extraction, pattern recognition and classification, which share in the development of BCI technology.

A. *Feature Extraction and Signal Processing*

The goal of feature extraction is removing noise and unnecessary information from raw signals while retaining important information to discriminate different classes of signals, moreover, reducing classification computational cost by reducing signal dimension. Neurophysiological knowledge can aid to decide brain signal feature that expected to hold most discriminative information for certain paradigm. Brain activity variations can be found in time, frequency and space domains depending on type of signal and its characteristics. So, features are extracted from signals by signal processing methods which reduce variability and yield similar presentation for all signals to be classified.

EEG signals processed as follows to retain discriminative features. First, is extracting time domain features which are related to changes in signal amplitude that occur during stimuli presentation time. The significant segment in the signal is that occurs after the intensification of a row/column hereafter called Post-Intensification segment. Since EEG signal provided is recorded as one signal and not segmented. Therefore, the first step is to extract Post-Intensification segments corresponding to stimuli presentation. As P300 appears after about 300ms of the stimulus, segments have been extracted by applying time window within 0-650ms after the beginning of row/column intensification. This window is considered large enough to capture the required features for efficient classification. Second step, in order to separate these signals from background activity and noise, band-pass filtering is applied. The raw signals are originally band-pass filtered at 0.1-60 Hz, extra filtering applied using band-pass filters with the following frequency ranges: 2-8, 0.1-10, 0.1-20, 0.1-30, 0.1-40Hz. These ranges chosen as cognitive activity very rarely occurs outside the range from 3-40Hz [9]. The winners of BCI competition II [4] [12] and competition III [13] used 0.5-30Hz and 0.1-20Hz band-pass filter ranges respectively. Filtering is followed by down-sampling. This time domain related approach allows removing unimportant information from high frequency bands, in addition to signal dimensionality reduction.

Secondly, is extracting spatial domain features. Raw data is not univariate time series (i.e. from one electrode), its acquired from 64 electrodes. Therefore, features extracted from several electrodes have to be combined in an efficient way. As changes in P300 peaks do not occur uniformly at all electrodes but are usually stronger over task cognitive brain regions. Algorithms automatically selecting an optimal electrode subset as done by [5] [13] and [11] could be used. Although reducing the number of electrodes could help in reducing data dimensionality, but reduced electrodes have to be adaptively selected with respect to each subject. In this work, this was applied using ICA or PCA. ICA is used as a blind source separation (BBS) approach. The signals measured on scalp are mixture of signals from independent sources in the cortex, deeper brain structures and noise. ICA is used to separate multi-channel EEG into several components, corresponding to sources in the brain or noise. By retaining only components having a P300-like spatial distribution or show P300-like waveforms, the signal-to-noise ratio can be improved. Consequently, classification can be performed with improved accuracy. The number of retained electrodes tested here is 32 electrodes. PCA is applied to rank the electrodes according to their importance. Accordingly, number of electrodes can be reduced by eliminating the least ranked ones. The number of retained electrodes tested here is 56, 48, 32, 22 and 16 electrodes.

After processing & feature extraction Post-Intensification segments transformed into feature vectors by concatenating samples from all retained electrodes vertically, normalized to zero mean and unit variance. Same processing steps applied on testing dataset before classification.

B. Pattern Recognition and Classification.

Classification problem is divided into two tasks. The first, is binary classification for each feature vector determining whether it contains P300 or not, therefore, the corresponding row/column is target or not-target. The second, is dealing with 36-characters classification, scores should be aggregated over 15 sequences to predict target character.

Non-parametric supervised learning classifiers were employed which depend on calculating distance; K-Nearest Neighborhood (KNN) and Euclidean Distance (ED). Parametric classifiers selection was guided by two approaches. First approach linear classifiers were applied as they never found to perform worse than non-linear classifiers [10]; Linear Discriminant Analysis (LDA), Fisher Linear Discriminant Analysis (FLDA), Linear Support Vector Machines (LSVM) and Generalized Anderson's Task (GAT) [14]. Second approach depends on extending the functionality of linear classifiers by regularization and combining multiple classifiers; Bayesian Linear Discriminant Analysis (BLDA) and combination of multiple Support Vector Machine (SVM) classifiers. BLDA is an extension of FLDA [6] [11]. Combination of multiple SVM is an extension of LSVM [12]. The outputs of multiple SVM classifiers are fused together in order to produce a single predicted character using two procedures. The first procedure hereafter called MSVM I, signals from each row/column averaged over sequences [4]. Second procedure hereafter called MSVM II, double averaging applied over sequences and multiple classifiers scores [12].

III. RESULTS AND DISCUSSION

Table I shows subjects 'A' and 'B' performances during 5th and 15th sequences. Performance is evaluated based on percentage of correctly predicted characters on test sets. *Table II* shows subject 'C' results provided as number of misspelled characters among sequences. Highest results among chosen frequency ranges (high cut-off frequency of the range are highlighted in Hz in both tables.

The results for extended linear classifiers without electrode reduction were ranked for BLDA (64) as the best performance followed by MSVM II (64) and MSVM I (64) at 5th and 15th sequences of 'A' and 'B'. For 'C' the same ranking holds for the three algorithms. All extended linear classifiers outperformed other linear classifiers. BLDA, Bayesian version of FLDA outperformed plain FLDA. As well as MSVM I and MSVM II outperformed LSVM. While for MSVM I and MSVM II as expected the double averaging method MSVM II outperformed the single averaging method MSVM I as shown in the results of *Table II*. BLDA outperformed all classifiers due to its robust and quick parameter estimation as discussed in [14].

For linear classifiers reasonable performance was achieved at the 15th sequence for GTA (64), LSVM (64), FLDA (64) and LDA (64) of 'A' and 'B'. For 'C', the ranking of algorithms differs than 'A' and 'B'. LDA (64) yields the best results, followed by FLDA (64) and GAT (64). LSVM (64) was the least of linear classifiers.

Concerning non-parametric classifiers ED and KNN both yielded the worst results compared to others. In addition, KNN is the only nonlinear classifier, its results support the approach of selecting linear classifiers followed in this work.

Applying ICA or PCA removes noise and artifacts, moreover, the great advantage of reducing the number of electrodes. As shown in *Table I* reducing the number of electrodes to half (32) improved performance of linear classifiers as observed with GTA, LSVM, and FLDA for 'A' & 'B' at 15th sequence. While for LDA, the performance degraded as LDA looks for linear combinations of variables which best explain the data similar to PCA [15] which means that reducing electrodes led to losing significant data. Similarly, results stand for 'C' as shown in *Table II.*

For extended linear classifiers electrode reduction to 32 and 48 electrodes almost provide same results compared to performance without electrode reduction for all 3 subjects.

Reducing number of electrodes to 16 or 22 degraded the performance of all classifiers with all 3 subjects. While reducing electrodes to 32 yielded almost same results for ICA or PCA alternatively, with all classifiers and subjects.

The performance was not only affected by combination of different feature extraction and classification algorithms but also affected by signal processing and subject performance.

Concerning signal processing phase, in most previous BCI trials, same filtering frequency range was applied to all subjects. Yet, in this work multiple band-pass filtering frequency ranges were applied which helped in reaching better performance than sticking to same range for all classifiers with all subjects. And so the results achieved in this work outperformed the results of competition winner although employing same classifier MSVM II [14].

Nevertheless, subject performance affected overall system performance; due to high subject performance variance, and variability between different subjects. This can be observed in performance of ‘A’ compared to ‘B’. Subject ‘B’ outperforms ‘A’ with all classifiers in 5th sequence. But due to the robust classification algorithms and score aggregation over sequences almost performance of both is the same at 15th sequence, supporting that for stimulus-driven BCI system the load of training is on BCI system not the subject.

IV. CONCLUSION

An efficient P300-based BCI system is discussed. Special emphasis on development of supervised MLAs discriminating EEG segments containing P300. Due to the use of P300 training load is on system and not subject.

TABLE I. RESULTS OF SUBJECTS “A” & “B” COMPETETION III 2004

Algorithm & Electrodes	5th Sequence			15th Sequence			Frequency	
	A	*B*	*Avg*	*A*	*B*	*Avg*	*A*	*B*
	%			*%*			*Hz*	
BLDA 64	69	77	73	98	98	98	40	10
BLDA 48	63	78	70.5	98	96	97		
BLDA 32	64	78	71	98	97	97.5		
MSVM II 64	70	75	72.5	98	96	97	30	10
MSVM II 48	65	74	69.5	97	95	96		
MSVM II 32	63	76	69.5	98	96	97		
MSVM I 64	49	76	62.5	95	96	95.5	30	10
MSVM I 48	54	70	62	95	95	95		
MSVM I 32	47	70	58.5	91	96	93.5		
GAT 64	47	63	55	89	91	90	20	10
GAT 48	44	61	52.5	93	96	94.5		
GAT 32	49	58	53.5	88	91	89.5		
LSVM 64	52	66	59	89	87	88	40	10
LSVM 48	50	54	52	90	88	89		
LSVM 32	54	70	62	91	93	92		
FLDA 64	49	56	52.5	84	90	87	20	10
FLDA 48	51	61	56	91	92	91.5		
FLDA 32	46	56	51	88	91	89.5		
LDA 64	29	67	48	75	91	83	30	10
LDA 48	28	66	47	65	90	77.5		
LDA 32	21	61	41	64	84	74		
ED 64	12	9	10.5	38	18	28	40	30
K-NN 64	5	4	4.5	8	7	7.5	10	40

TABLE II. RESULTS OF SUBJECT “C” COMPETETION II 2003

Algorithm & Electrodes	Number of Sequences										Frequency
	Number of misspelled characters										*High cut-off*
	1	*2*	*3*	*4*	*5*	*6*	*7*	*9*	*11*	*15*	*Hz*
BLDA 64	5	2	0	0	0	0	0	0	0	0	40
BLDA 32	4	2	1	1	0	0	0	0	0	0	
MSVM II 64	7	5	3	0	0	0	0	0	0	0	20
MSVM II 32	7	2	2	1	0	0	0	0	0	0	
MSVM I 64	9	6	3	1	0	0	0	0	0	0	20
MSVM I 32	12	5	4	1	0	0	0	0	0	0	
LDA 64	17	15	9	4	3	0	0	0	0	0	20
LDA 32	13	7	4	3	1	1	0	0	0	0	
FLDA 64	22	10	7	4	4	4	2	0	0	0	40
FLDA 32	18	9	5	4	1	1	0	0	0	0	
GAT 64	20	16	9	8	6	5	4	0	0	0	30
GAT 32	22	11	9	5	2	0	0	0	0	0	
LSVM 64	17	16	11	10	6	8	5	2	3	1	10
LSVM 32	23	16	15	7	5	7	5	0	0	0	
ED 64	25	25	24	25	24	22	23	17	16	13	10
K-NN 64	27	28	25	23	21	19	18	20	16	12	10

Different approaches were demonstrated to address the variance in subject performance and variance between different subjects. Overcoming the challenging classification problem due to EEG signal nature was accomplished through applying time domain and spatial domain feature extraction. Applying ICA and PCA as spatial feature extraction helped to reduce number of electrodes to half its original size, and also improved performance which consequently, reduced classification time. This achievement could help in using P300 Speller BCI system beyond proof-of-concept and allow their widespread application out of laboratories to real life.

REFERENCES

[1] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw, “BCI2000: a general-purpose brain-computer interface (BCI) system,” IEEE Transactions on Biomedical Engineering, vol. 51, no. 6, pp. 1034–1043, June 2004.

[2] B. Blankertz, K.-R. Mueller, G. Curio, T. Vaughan, G. Schalk, J. Wolpaw, A. Schloegl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schroeder, and N. Birbaumer, “The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single trials,” IEEE Transactions Biomedical Engineering, vol. 51, no. 6, pp. 1044–1051, June 2004.

[3] B. Blankertz, K.-R. Müeller, D. Krusienski, G. Schalk, J. Wolpaw, A. Schlögl, G. Pfurtscheller, J. del R. Millan, M. Schröder, and N. Birbaumer, "The BCI competition iii: Validating alternative approaches to actual BCI problems," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 14, no. 2, pp. 153–159, June 2006.

[4] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, “BCI competition 2003 - dataset IIb: Support vector machines for the P300 speller paradigm,” IEEE Transactions on Biomedical Engineering, vol. 51, no. 6, pp. 1073– 1076, June 2004.

[5] A. Rakotomamonjy, V. Guigue, G. Mallet, and V. Al-varado, "Ensemble of SVMs for improving brain-computer interface P300 speller performances.," International Conference on Artificial Neural Networks (ICANN)., Springer Berlin / Heidelberg, vol. 3696 pp. 45-50, November 2005.

[6] U. Hohffmann, G. Garcia, J.-M. Vesin, and T. Ebrahimi, "Application of the evidence framework to brain-computer interfaces" IEEE Engineering in Medicine and Biology Conference., vol. 1, pp. 446 – 449, September 2004.

[7] B. Blankertz, BCI Competition III Webpage. [Online]. Available: http://ida.first.fraunhofer.de/projects/bci/competition_iii

[8] B. Allison,"P3 or not P3: Toward a Better P300 BCI", ph.D. dissertation, University of California, San Diego 2003.

[9] The BCI2000 Project. Facilitating brain-computer interface research and applications. Third BCI2000 Workshop, Rome, December 5-6, 2007. Available: http://www.bci2000.org

[10] F. Vojtěch, H. Václáv "Statistical Pattern Recognition Toolbox for Matlab User’s guide," June 24, 2004.

[11] U. Hoffmann, J-M Vesin, T. Ebrahimi, and K. Diserens, " An efficient P300-based brain-computer interface for disabled subjects ", Journal of Neuroscience Methods, vol. 167, no. 1, pp. 115-125, 2008.

[12] A. Rakotomamonjy, and V. Guigue, " BCI Competition III: Dataset II- Ensemble of SVMs for BCI P300 Speller," IEEE Transactions on Biomedical Engineering, vol. 55, no. 3, pp. 1147 – 1154, March 2008.

[13] A. Rakotomamonjy, "Variable Selection Using SVM-based Criteria", Journal of Machine Learning Research, vol. 3, pp. 1357–1370, 2003.

[14] A. Selim, M. Abdel Wahed & Y. Kadah, “Machine Learning Methodologies in Brain Computer-Interface Systems,” Proceedings of 26th National Radio Science Conference, NRSC'2009, Future University, March 17-19, 2009.

C. M. Bishop, "Pattern Recognition and Machine Learning", 2006.