NTAV/SPA 2012
27-29TH SEPTEMBER, 2012, ŁÓDŹ, POLAND
NEW TRENDS IN AUDIO AND VIDEO / SIGNAL PROCESSING ALGORITHMS, ARCHITECTURES, ARRANGEMENTS AND APPLICATIONS

Cluster Analysis of Canonical Correlation Coefficients for the SSVEP Based Brain-Computer Interfaces

Paweł Poryzała
Institute of Electronics,
Technical University of Lodz
Lodz, Poland
e-mail: poryzala@p.lodz.pl

Andrzej Materka
Institute of Electronics
Technical University of Lodz
Lodz, Poland
e-mail: materka@p.lodz.pl

Marcin Byczuk
Institute of Electronics
Technical University of Lodz
Lodz, Poland
e-mail: byczuk@p.lodz.pl

Abstract — **In this paper, a novel method for detecting steadystate visual evoked potentials (SSVEP) using multiple channel electroencephalogram (EEG) data is presented. Accurate asynchronous detection, high speed and high information transfer rate can be achieved after a short calibration session. Spatial filtering based on the Canonical Corelation Analysis method proposed in [1] is used for identifying optimal combinations of electrode signals that cancel strong interference signals in the EEG. Data from a test group consisting of 21 subjects are used to evaluate the new methods and to compare results to standard spectrum analysis approach. Conducted research, for different length signal segments and five visual frequencies, showed improvement of both classification accuracy and detection speed.**

Index Terms — Brain-Computer Interface (BCI), Electroencephalogram (EEG), Steady-State Visual Evoked Potential (SSVEP) detection

I. Introduction

Studies on the development of the Brain-Computer Interfaces (communication systems, that do not depend on the brain's normal output pathways of peripheral nerves and muscles [2]) have more than 20-year history. BCI devices may allow people with disabilities, including paralysed people, use the computer and other technical equipment, on a par with other users. Over the years, most widely represented group of devices are non-invasive BCI systems with electroencephalographic (EEG) brain activity monitoring.

At the moment, the most commonly used EEG-based BCI systems employ event-related synchronization of μ and β rhythms (ERD/ERS), event-related potentials (ERP) and steady-state visual evoked potentials (SSVEP). Information transfer rate (ITR, introduced in [3]) is used by majority of the BCI laboratories and research groups to evaluate BCI system performances. This measure depends on three factors: speed, accuracy and number of targets. It is proved, that currently the SSVEP approach provides the fastest and the most reliable communication paradigm for the implementation of a non-invasive BCI system [4].

High speed and accuracy, sufficient number of targets for a particular task are essential for BCI system in order to become a practical device. Today a number of signal processing methods for detection and extraction of SSVEPs exist. From simple methods for detecting a single frequency component in a single electrode signal [5], through most widely used spectrum analysis methods [6], [7] up to multichannel spatial filtering and detection methods [8], [1].

In this paper, a novel approach for multichannel detection of SSVEP responses is proposed. System, after a simple calibration session, is able to work asynchronously with improved (in relation to spectrum analysis method) detection speed and accuracy (thus higher ITR).

The paper is organized as follows. The second section discusses the details of the proposed method. Off-line experiment conducted to prove the algorithm quality are presented in the third session. Fourth section contains results and discussion. Conclusions are presented in the last section.

II. Detection Method

In this section, the proposed Cluster Analysis Canonical Correlation (CACC) method for detection of SSVEPs is discussed. It is based on the coefficients derived from the Canonical Correlation Analysis (CCA) which is described in what follows.

A. Canonical Correlation Analysis

CCA method is used for finding the correlations between two sets of multi-dimensional variables. It was first used for SSVEP detection in [1] and was further developed in [9].

CCA method seeks for a pair of linear combinations $\mathbf{w}$ and $\mathbf{v}$, for two sets of data $\mathbf{Y}$ and $\mathbf{X}$, such that the correlation

$$\rho_1 = \mathrm{cor}(\mathbf{S}, \mathbf{U}) \tag{1}$$

between the first pair of canonical variables $\mathbf{S} = \mathbf{w}^T\mathbf{Y}$ and $\mathbf{U} = \mathbf{v}^T\mathbf{X}$ is maximized. Consecutive pairs of linear combinations, canonical variables and canonical correlation coefficients can be obtained, but the maximum number of pairs equals the number of variables in the smallest of two sets ($\mathbf{Y}$ and $\mathbf{X}$).

As far as the CCA method is used for SSVEP detection: $\mathbf{Y}$ refers to the set of N_y multi-channel EEG signals and $\mathbf{X}$

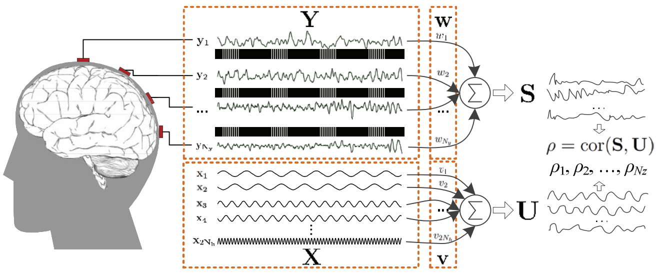


Fig. 1: An illustration for usage of the CCA method in EEG signal analysis. Matrix $\mathbf{Y}$ is where data from N_y EEG channels is stored. $\mathbf{X}$ is an ideal, reference SSVEP response, containing both sinus and cosinus components for N_h harmonics.

refers to the set of $2N_h$ reference signals (Fig. 1). In the rows of the $\mathbf{X}$ reference matrix, the sinus and cosinus components for all N_h harmonics of the stimulation frequency are stored:

$$\mathbf{X} = \begin{bmatrix} \sin(2\pi f t) \\ \cos(2\pi f t) \\ \dots \\ \sin(2\pi N_h f t) \\ \cos(2\pi N_h f t) \end{bmatrix}. \tag{2}$$

CCA finds the maximum canonical correlation with respect to weight vectors $\mathbf{w}$ and $\mathbf{v}$ by solving the following problem:

$$\begin{aligned} \max_{\mathbf{w},\mathbf{v}} \rho &= \frac{\mathrm{cov}[\mathbf{S},\mathbf{U}]}{\sqrt{\mathrm{var}[\mathbf{S}]\,\mathrm{var}[\mathbf{U}]}} \\ &= \frac{E[\mathbf{SU}]}{\sqrt{E[\mathbf{S}^2]E[\mathbf{U}^2]}} \\ &= \frac{E[\mathbf{w}^T\mathbf{Y}\mathbf{X}^T\mathbf{v}]}{\sqrt{E[\mathbf{w}^T\mathbf{Y}\mathbf{Y}^T\mathbf{w}]E[\mathbf{v}^T\mathbf{X}\mathbf{X}^T\mathbf{v}]}}. \end{aligned} \tag{3}$$

When CCA is used in frequency recognition of the SSVEP-based BCI system, where there are N_f targets (stimulus frequencies $f_1, f_2, \dots, f_{N_f}$), the same number of reference matrices must be used (Fig. 2).

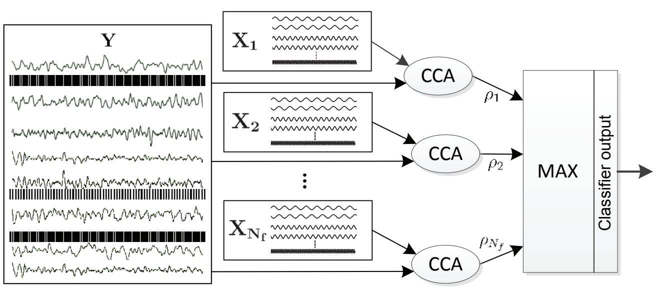


Fig. 2: An illustration for usage of the CCA method in different frequency components recognition of the SSVEP-based BCI where there are N_f targets. $\mathbf{X_i}$ is the response reference matrix for the i-th stimulus frequency.

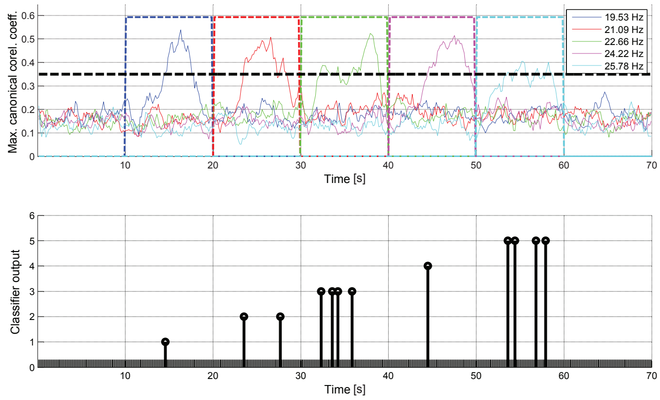
For each pair of multi-channel EEG and reference signals, a maximum canonical correlation coefficient is obtained and it can be used for frequency recognition. As proposed in [9] user's command is recognized as

$$C = \max_i \rho_i, \quad i = 1, 2, \dots, N_f, \tag{4}$$

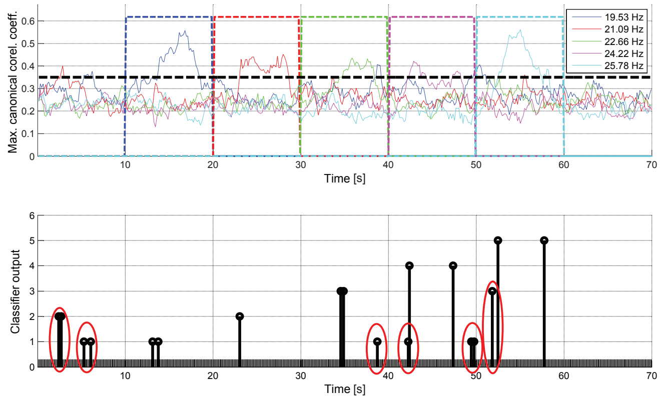
where ρ_i is the CCA coefficient obtained with the reference signal frequency being $f_1, f_2, \dots, f_{N_h}$.

B. Encountered CCA problems

Original CCA method, even in conjunction with thresholding of maximum canonical correlation coefficients for each stimulus frequency, does not seem reliable in practical, asynchronous SSVEP BCI system. Main problem is related to strong dependence of measured EEG signals against user psychophysical state (Fig. 3). This state changes with the ongoing measurement session and between different days (when user eg. did not sleep well). In such cases, the background, non-stimulated EEG activity is increased. Brief moment of relaxation often improves recorded signal quality, but this is usually only a short-term effect.



(a) good psychophysical state



(b) bad psychophysical state

Fig. 3: Comparison of the classification results of the relevant parts of the signal for the AL1 user in two test sessions. In the second case wrong system decisions were marked red.

In Fig. 3b it is clearly visible that all of the canonical correlation coefficients have greater variability over time (often reaching established threshold value, resulting in false detections). In this particular example SSVEP BCI system is not able to distinguish between working and idle state classes properly. There is also only a little margin to rise threshold value due to the low canonical correlation coefficient values in segments which involved stimulation.

C. Cluster Analysis Canonical Correlation

Original CCA method uses a single canonical correlation coefficient (with the highest value) for each of the N_f SSVEP response patterns. CACC method uses three highest valued correlation coefficients as features. Detection and idle states can be accurately identified with k-means cluster analysis performed separately in each of the feature spaces (Fig. 4).

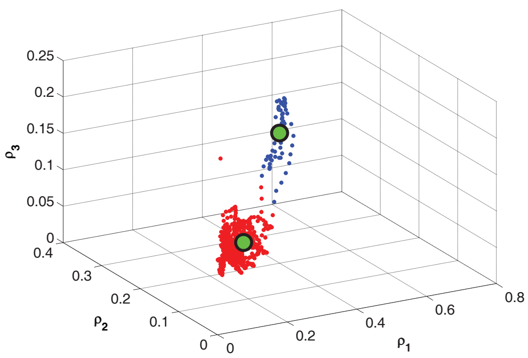


Fig. 4: Sample result of the k-means cluster analysis in the correlation coefficients feature space. Idle class was marked red, detection class was marked blue. Centroids of both classes were marked green.

Distance which can be measured between centroids of both detection and idle classes in feature spaces for each stimulus frequency, varies between the subject and frequency used for stimulation. Its value must be determined during the training phase, therefore BCI system work should be divided into two stages:

1) calibration session: At this stage (Fig. 5) the objective of the algorithm is to identify the distances between centroids of detection and idle classes. This value is characteristic for each of the frequencies used for stimulation. The user is instructed to move his/her eyes (but not faster than every 5 seconds) between all stimulation symbols.

In the first step, a set of response patterns for each of the stimulation frequencies used ($\mathbf{X_i}, i = 1, 2, ..., N_f$) is built. As a result of canonical correlation of N_y EEG source channels in the detector window $\mathbf{Y}$ sequentially with patterns $\mathbf{X_i}$, one gets sets of three factors: ρ_{1i}, ρ_{2i} and ρ_{3i}. Each of the sets can be represented as a point ρ_i in the the feature space constructed on the basis of canonical correlation coefficients of the source EEG data with the i-th response pattern.

Along with each successive point ρ_i in particular feature space, k-means cluster analysis is performed and mutual

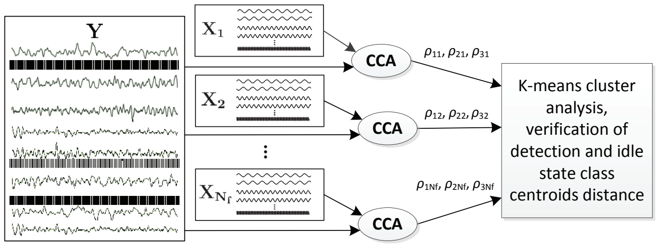


Fig. 5: System in calibration mode.

distance between two classes (detection and idle state) is examined. Euclidean metric is used:

$$d(B_i, D_i) = \sqrt{\sum_{j=1}^{3} (\rho_{B,ji} - \rho_{D,ji})^2}, \tag{5}$$

where B_i and D_i denote the points where the idle and detection class centroids lay in the i-th feature space. Calibration of the frequency f_i ends when the B_i and D_i centroid distance is large enough:

$$d(B_i, D_i) \geq \beta \tag{6}$$

and after adding e.g. the last 25 points to appropriate feature space, the distance was not changed by more than 10%.

Based on the analysis of recorded EEG data and our practical investigations, $\beta = 0.25$. Its value is a compromise between the accuracy (especially for lower quality signals) and the time of detection. Too high β value results in increased number of false negative errors, and too small increases false positives.

The training session ends upon completion of the calibration for all N_f frequencies. If the calibration procedure lasts over one minute, the system reports a problem with particular frequency.

2) working mode: This is the target operating mode, in which device is used for communication (Fig. 6). All calibrated data (locations of the detection and idle class centroids in each of N_f feature spaces) are used to improve classification at this stage.

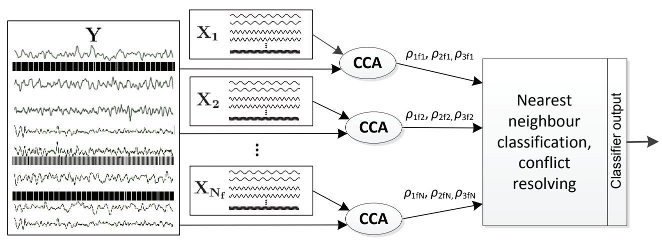


Fig. 6: System in working mode.

Like in the calibration mode, as a result of canonical correlation of the EEG source data ($\mathbf{Y}$) with subsequent response patterns X_i, sets of three coefficients: ρ_{1i}, ρ_{2i} and ρ_{3i} (a point ρ_i in a three dimensional feature space) are obtained. Each point is classified (nearest neighbours method) to one of the classes B_i or D_i.

If, during the classification in each of N_f feature spaces, none or exactly one point ρ_i was classified to D_i class, system will detect respectively class zero (idle state) or number i of particular feature space. If more than one point, represented by the canonical correlation analysis coefficients of source data and response pattern X_i, will be qualified to the detection classes, a conflict occurs.

Conflict situations are solved by using the distance of each of the conflicted points ρ_i from the point laying on the line passing through centroids of both B_i and D_i classes, and lying half-way between them. The classifier output is determined as the number of the i-th feature space in which the distance was the greatest.

After successful detection of responses at any of the stimulus frequencies, all data in detector window $\mathbf{Y}$ are replaced with zeros. This prevents multiple detection of the same symbol. In addition, after each classification, 700 ms of the EEG data will not be utilized (classifier will not take any decisions). This will give the user of the BCI system time for gaze shifting.

III. Offline experiments

The experiments were carried out at the Institute of Electronics, Technical University of Lodz. Fig. 7 presents the layout of the measurement stand.

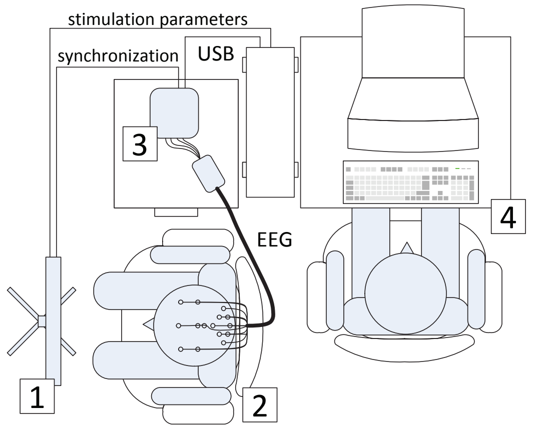


Fig. 7: Layout of the measurement stand: visual stimulator (1), subject (2), EEG recording device (3) and operator (4).

Subjects sat in the front of a visual stimulator (described in the next section) on a comfortable, ergonomic chair. Measurements were carried out in a room with a window on the south side, curtained with a light impermeable material blind and a standard fluorescent light switched on. Light conditions during all experiments were the same.

A. Subjects

Twenty one healthy subjects (ten women and eleven men, age range 16-33 years, with the average of 22.2 years and a standard deviation of 3.4 years) participated in this study. For each subject, two measurements were carried out on different days.

Four subjects previously used our BCI system. None of the subjects had neurological or visual disorders (glasses or contact lenses were worn where appropriate). Subjects did not receive any financial rewards.

In the early stages of the experiment, users were qualified to one of three groups:

1) Group A (best results, 5 subjects): Subjects who in most cases had earlier contact with the device (in our previous studies and tests).

2) Group B (average results, 11 subjects): The most widely represented group. Subjects who were not familiar with the idea of a BCI device, but actively participated in the experiments.

3) Group C (poor results, 6 subjects): Subjects with concentration problems or very high unstimulated, spontaneous brain activity

This classification helped to investigate system parameters in relation to a specific group of users.

B. Visual Stimulator

A universal, computer driven LED stimulator was used for stimulation. Each stimulation symbol (Fig. 8) consisted of three LEDs: two stimulation lights with a diameter of 5mm positioned on the lower right and lower left quarter of the visual field of each eye retina and one fixation light with a diameter of 3mm placed in the center of visual field. Distance from visual stimulator to subject eyes was equal to 50 centimetres.

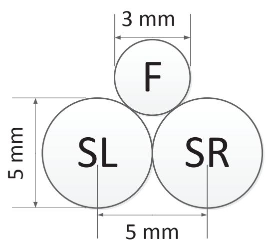


Fig. 8: A view of stimulating lights (SL, SR) and a fixation light (F) on the screen of stimulator.

Stimulation lights flash with the same frequency alternatively in phase (alternate half-field stimulation technique [10]). Fixation light is used for two purposes: the subject is expected to concentrate his/her sight on it; additionally it provides a feedback information about amplitudes of corresponding SSVEPs detected in the subject EEG signal.

Visual stimulator had five sets of LEDs forming stimulation symbols in five different colors (each set had stimulation and fixation LEDs of the same color): white, blue, green, yellow and red. Luminous intensity of each LED used was approximately 1000mcd.

C. EEG Recording

Equipment from g.tec (Graz, Austria) was used for EEG measurements: g.USBamp biosignal amplifier, g.GAMMAbox active electrode driver and g.GAMMAcap with sixteen

Ag/AgCl active electrodes. Seven electrodes over the primary visual cortex (positions PO7, PO3, O1, OZ, O2, PO4 and PO8) and nine electrodes evenly distributed over the remaining cerebral cortex (positions P3, PZ, P4, C3, CZ, C4, F3, FZ and F4) were used for recording. A ground electrode was placed on CPZ position. A reference electrode was placed on right ear lobe (position A2). The EEG signals were bandpass filtered between 2.0-60.0 Hz with a notch filter for 50 Hz power line frequency suppression, amplified and sampled at 600 Hz.

EEG signals were recorded with a home-made software package - BioStudio [11] which was able to drive visual stimulator and processed measured signals in order to compute biofeedback information for stimulation symbols.

D. Experimental paradigm

Subjects were instructed to focus their gaze on fixation LED and flickering lights below it to produce SSVEPs. Each measurement lasted for several minutes and consisted of five stimulus sequences (one sequence for each color, only one stimulation symbol switched on at a time). The first sequence began a few seconds after starting the measurement (time required for stabilization of electrode-skin connection impedance and possible adjustments of subjects' position on the chair to reduce the EMG signals). Stimulation frequencies were chosen to match the discrete Fourier transform frequencies used in the subsequent analysis (in order to minimize spectral leakage). Each sequence contained 27 different stimulation frequencies in the range of about 7–47 Hz.

Each stimulation lasted eight seconds, followed by a 2-second pause before the next stimulation (Fig. 9). Additionally a brief pause followed each sequence (several up to tens of seconds). This pause was intended for position adjustments on the chair and subject relax with eyes closed (EEG signal was still being recorded).

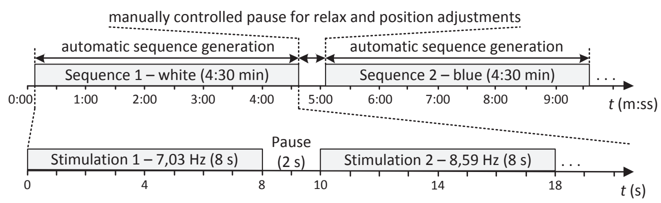


Fig. 9: Timing of each trial.

Binary signal from visual stimulator indicating stimulation state (on/off) was recorded along with the subject EEG signal from all sixteen channels.

The original EEG data for subjects were re-sampled ($F_s = 200$ Hz) and divided into shorter fragments, containing several stimulation patterns. Algorithm was tested with window lengths of 1.28, 2.56 and 5.12 seconds and data window moved with a step of 0.16 s.

Results of the proposed method were compared to standard spectrum analysis SSVEP detection approach: power spectral density of EEG signal was computed in the sliding window (frequency resolution of about 0.78 Hz). For each predefined discrete frequency of stimulation a signal to background ratio (SBR) was estimated [12]. The frequency of the maximum SBR, after it was compared with the threshold value, was decided to be the intended target of the user. This algorithm was executed for all possible bipolar source electrode combinations:

$$C^2_{N_y} = \binom{N_y}{2} = \frac{N_y!}{2!(N_y-2)!} \tag{7}$$

in order to find Best Bipolar Combination (BBC). In analysed case ($N_y = 16$) 120 bipolar channels had to be processed.

TABLE I: Results in Group A

Window Length [s]	accuracy [%]		det. speed [s]		ITR [bpm]	
	BBC	CACC	BBC	CACC	BBC	CACC
1.28	91.19	90.27	2.55	2.28	40.38	43.82
2.56	94.12	93.05	2.47	2.52	45.74	43.29
5.12	91.53	94.88	4.02	3.35	25.85	34.52

IV. Results and discussion

Binary markers (stimulation on and off events) stored in parallel with the EEG data and the known stimulation sequence for each color were used to verify performance of the proposed detection algorithm. Classification results for each user were assessed in terms of accuracy, average detection time and the information transfer rate and were afterwards averaged in each of the subject groups.

A. Group A

High accuracy of both SSVEP detection methods is proved (Table I). The increase in detection accuracy with the increase of window length is negligible. Measured mean detection times increase as the window length is extended (this is a known problem an can be easily solved in practical system by use of multiple, different length parallel detectors). Information transfer rates are similar in case of both algorithms.

B. Group B

The biggest increase of accuracy of the CACC method over the BBC algorithm was observed in this group (Table II). Depending on the window length, it was from 7 up to 11%. There is also a noticeable rise of detection accuracy with the increase of window length. As in the previous group, average detection times are similar (particularly for shorter windows), but detection usually took about 0.8–1.0 second longer. As far as the information transfer rate is considered, CACC method seams to be clearly better than BBC because of both: shorter detection times and higher accuracy.

C. Group C

Increase of detection accuracy for the CACC method over the BBC algorithm in this group was from 5 up to 8% (Table III). Similarly to the first group, the increase in detection accuracy with the increase of window length is negligible. As far as the average detection speed is considered, BBC method

TABLE II: Results in Group B

Window Length [s]	accuracy [%]		det. speed [s]		ITR [bpm]	
	BBC	CACC	BBC	CACC	BBC	CACC
1.28	63.15	70.74	3.39	3.02	11.24	17.18
2.56	68.26	79.65	3.26	3.15	14.46	22.58
5.12	68.33	78.51	5.06	4.15	9.34	16.51

TABLE III: Results in Group C

Window Length [s]	accuracy [%]		det. speed [s]		ITR [bpm]	
	BBC	CACC	BBC	CACC	BBC	CACC
1.28	45.07	50.98	4.02	5.06	3.44	2.14
2.56	47.23	53.22	4.75	5.12	3.39	4.06
5.12	47.12	55.17	5.35	5.72	2.99	4.54

is faster (difference of about 1 s for the shortest window and about 0.4 s in remaining cases).

V. Conclusions

Results clearly show that research on multichannel detection methods are important and can significantly improve classification accuracy, detection times and overall communication speed. The proposed detection method improves the classification accuracy in the groups of subjects with the average (Group B) and poor (Group C) results. In the group of users with the best results (Group A), there was no clear improvement of the SSVEP detection accuracy. Average detection times for both algorithms are similar in most cases (but there were differences of up to 1 second). Information transfer rate in many cases (especially for Groups B and C) was higher for the CACC method, which is due to greater classification accuracy of this method. What is important, only a short off-line calibration session was necessary to achieve such results.

At the moment many of the BCI systems are at the stage of laboratory demonstrations. This is mainly due to high user variation, BCI illiteracy phenomenon and low communication speeds (low ITR). New spatial filtering and detection methods will make it possible to overcome this limitations. In the presented research, each of 21 subjects was able to communicate in the off-line experiments and 16 subjects (Groups A and B) reached substantial information transfer rates. These results encourage further development of the proposed detection method and its implementation in the online BCI system, what will be the subject of our future work.

Acknowledgment

This work is partially supported by Polish Ministry for Science and Higher Education grant NN515 520838.

References

[1] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, Jun. 2007.

[2] J. Wolpaw, N. Birbaumer, W. Heetderks, D. McFarland, P. Peckham, G. Schalk, E. Donchin, L. Quatrano, C. Robinson, and T. Vaughan, "Brain-computer interface technology: a review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, Jun. 2000.

[3] J. Wolpaw, H. Ramoser, D. McFarland, and G. Pfurtscheller, "EEG-based communication: improved accuracy by response verification," *IEEE Trans. Rehabil. Eng.*, vol. 6, no. 3, pp. 326–333, Sep. 1998.

[4] Y. Wang, X. Gao, B. Hong, C. Jia, and S. Gao, "Brain-Computer Interfaces Based on Visual Evoked Potentials," *IEEE Eng. Med. Biol. Mag.*, vol. 27, no. 5, pp. 64–71, Sep. 2008.

[5] J. S. Victor and J. Mast, "A new statistic for steady-state evoked potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 78, no. 5, pp. 378–388, May 1991.

[6] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical VEP-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 234–240, Jun. 2006.

[7] A. Materka, M. Byczuk, and P. Poryzala, "A Virtual Keypad Based on Alternate Half-Field Stimulated Visual Evoked Potentials," in *Int. Symp. on Inf. Technol. Converg., ISITC 2007*, 2007, pp. 296–300.

[8] O. Friman, I. Volosyak, and A. Graser, "Multiple Channel Detection of Steady-State Visual Evoked Potentials for Brain-Computer Interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 742–750, Apr. 2007.

[9] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *Journal of Neural Engineering*, vol. 6, no. 4, p. 046002, Jun. 2009.

[10] A. Materka and M. Byczuk, "Alternate half-field stimulation technique for SSVEP-based brain-computer interfaces," *IET Electronic Letters*, vol. 42, no. 6, pp. 321–322, Mar. 2006.

[11] P. Poryzala, M. Byczuk, and A. Materka, "Modular, plugin-based computer software for on-line analysis of electroencephalographic data in Brain-Computer Interfaces," in *II Forum Innowacji Mlodych Badaczy*, Nov. 2011.

[12] M. Byczuk, P. Poryzala, and A. Materka, "On possibility of stimulus parameter selection for SSVEP-based brain-computer interface," in *Man-Machine Interactions 2 (Advances in Intelligent and Soft Computing)*, ser. Advances in Intelligent and Soft Computing, T. Czachorski, S. Kozielski, and U. Stanczyk, Eds. Springer, 2011, vol. 103, pp. 57–64.