Fast BCI Calibration
Comparing Methods to Adapt BCI Systems for New Subjects

Jean Thorey[1], Parvaneh Adibpour[1], Yohei Tomita[1], Antoine Gaume[1],
Hovagim Bakardjian[1,2,3], Gérard Dreyfus[1] and François-B. Vialatte[1,2]

[1]*SIGMA Laboratory, ESPCI ParisTech, Paris, France*

[2]*LABSP, Riken BSI, Wako-Shi, Japan*

[3]*IHU-A-ICM, Pitié-Salpêtrière Hospital, Paris, France*

Keywords: Brain Computer Interface, EEG, SSVEP, Calibration.

Abstract: A Brain Computer Interface (BCI) is a system where a direct connection is established between the brain and a computer, providing a subject with a new communication channel. Unfortunately, BCI have many drawbacks: signal recording is problematic, brain signatures are non reproducible from individual to individual, etc. A dependent-BCI prototype, the BrainPC project, was developed in the SIGMA laboratory. Electroencephalographic (EEG) signals collected by a BrainAmp amplifier in responses to flickering light stimuli (Steady State Visual Evoked Potentials) are converted into machine-readable commands. This system is coupled with a human-machine interface. We propose a solution for fast calibration of the automatic detection of SSVEP between experimental subjects. We tested different calibration methods; harmonic and electrode selections were shown to be the most efficient methods.

1 INTRODUCTION

Brain–Computer Interfaces (BCI) are communication systems that enable users to send commands to a computer by using only their brain activity (Nicolelis, 2011). This activity is generally being measured through EEG, which is a noninvasive technique for recording brain electrical activity at the surface of the scalp. In a BCI, the brain signals are recorded and analyzed to extract features that represent the messages buried inside the EEGs. Then a translation algorithm is needed to convert the features to a command which is supposed to be sent to the computer or external machine. It is through this procedure that disabled people can control a computer screen or navigate a wheelchair (Wolpaw et al., 2002). SSVEP-based BCIs are those BCIs that allow the users to communicate with a computer or machine, by SSVEP responses that are generated in their brain by looking at a repetitive visual stimulus. Steady State Visual Evoked Potential (SSVEP) is an oscillatory activity in human visual cortex that is phase locked to repetitive visual stimulation (Vialatte et al., 2010). Studies on developing SSVEP-based BCIs have used several algorithms for detecting the SSVEPs. Most of the studies in literature identify user's intended target by calculating the frequency spectrum analysis of the signal and this is typically implemented using the Fourier Transform, particularly Fast Fourier Transform (FFT). Detection is usually based on considering a threshold for the power spectrum at stimuli frequencies (Vialatte et al., 2010). There are also other recent studies, introducing new methods for detection of SSVEPs. (Friman et al., 2007, Lin et al., 2006, Bin et al., 2009, Zhang et al., 2012)

For a BCI to have applicability in daily life, it is very important to make it work in different situations and for different subjects. This is not always easy due to the subject variability in the spatial patterns and spectrotemporal characteristics of brain signals (Volosyak et al., 2010). This subject variability makes the pattern recognition part quite difficult. For solving this issue, a rather long calibration phase is usually added to BCI experiments in order to collect EEG data from the subject to train the classifier or to adapt the stimuli parameters for each subject. The main problem with the calibration phase is the long times it takes for

Thorey J., Adibpour P., Tomita Y., Gaume A., Bakardjian H., Dreyfus G. and Vialatte F..
Fast BCI Calibration - Comparing Methods to Adapt BCI Systems for New Subjects.
DOI: 10.5220/0004180806630669
In *Proceedings of the 4th International Joint Conference on Computational Intelligence* (SSCN-2012), pages 663-669
ISBN: 978-989-8565-33-4

recoding the EEGs. For several applications (*e.g.* video games or neurorehabilitation), a fast calibration is necessary. Some previous studies have proposed methods for reducing the calibration time but most of them require a database of recordings from different subjects or several past recordings from the same BCI user. Krauledat et al. (2008) showed in their study on motor imagery BCI users, how predefined spatial filters and classifiers on the recorded data of previous training sessions of the same user would eliminate the need for a whole calibration phase at the beginning of each online experiment. To do this, they adjusted the bias of the classifier at the beginning of the online experiment. However, their good classification results were showing the power of their method for session to session transfer for the same subject but not for inter-subject variability. Lotte (2011) proposed a method based on generation of artificial EEG trials from the few previous collected trials in order to increase the training data set of classifier. Generating artificial EEG trials was based on segmentation of the data from different trials and then concatenating the segments from different trials to make new trials. Shishkin et al. (2011) proposed to use single stimulus for the calibration phase in a P300-based BCI in order to avoid the conflicts of non-target stimuli. The performances of their BCI system did not deteriorate significantly even when trained using a single-stimulus protocol. Wang et al. (2006) showed that user variability could be reduced by adapting channel, stimulus frequency and speed of command detection for each subject. Volosyak et al. (2010) compared two calibration methods of single LED and multi-target group LED stimuli for exploring the best stimulation frequencies. They found a strong correlation between the selected stimulation frequencies through both methods. They concluded they could shorten the calibration time significantly by using the multi-target group LED stimuli for detecting the best stimuli frequencies.

We investigate here methods for fast calibration of an SSVEP BCI based on selection of the channels and dominant frequency between the first and second harmonics of stimulation frequency independently for each subject. Such a method would allow us to use the system directly on new subjects, without long calibration times, but nevertheless exploiting previously collected data to design an optimal classifier.

2 METHODS

2.1 Experimental Paradigm

A virtual phone keypad was used as the interface, with 9 digits displayed, each of them were flickering with a pre-decied frequency (5.45 , 20, 8.75, 4.62, 6.67, 7.5, 12, 5 and 4 Hz). The background colour was black and the screen refresh rate was 60 Hz. This display was was realised using Cogent Graphics developed by John Romaya at the LON at the Wellcome Department of Imaging Neuroscience.

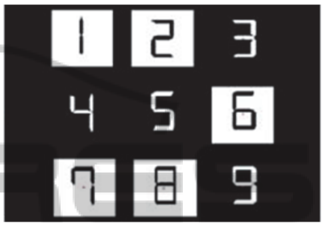

Figure 1: SSVEP stimulation interface.

2.2 Data Acquisition

EEGs were recorded with a BrainAmp amplifier; with a sampling rate of the 500 Hz. 16 active electrodes were placed over the head, according to the 10-20 international system for electrode placement. The electrodes covered the frontal, temporal and occipital sites (Fp_1, Fp_2, F_3, F_4, F_7, F_8, C_3, C_4, T_3, T_4, T_5, T_6, Po_1, Po_2, O_1, O_2), with the reference and grounds placed in the central positions (F_z and P_z).The subjects were told to relax and focus during 30 seconds on each command consecutively. 30 seconds of resting state eyes open were also recorded in front of a black screen at the beginning of each recording session.

We recorded 7 subjects. All subjects were young adults without any known history or actual brain disorder or anomaly.

2.3 Feature Extraction

The overall workflow of signal processing is as follows: (1) supervised feature extraction, for all subjects; (2) calibration, for each subject independently; (3) command classification and performance evaluation. After having selected a set of seven relevant features for the BCI system, we use the database collected to evaluate the performance of the BCI on new unknown subjects (classification is detailed in section 4). We compared the results obtained on the raw data, with results obtained after calibration of the data (calibration is explained in section 3). When applying calibration, it is applied for each subject: the “unknown” test subject as well as those used as a training reference for the classifier. All the selected features are pre-processed, for each subject, using these calibration techniques. We will now explain the supervised feature selection approach used.

Most BCIs are designed around a pattern recognition approach. In an EEG-based BCI, the first step is to extract features describing the relevant information buried in the EEG signals. They are then fed into a classifier which identifies the class which these features belong to. For detection of SSVEPs, we extracted the following features from the signals:

Fourier Peak: For detecting the SSVEPs in the signal, one first need to transform the time domain EEG into the frequency domain using Fourier transform. Once it is transformed into frequency domain, peaks at stimulation frequency and its harmonics are detectable. For detecting these peaks, we took the maximum amplitude in the Fourier spectrum of the signal at a small margin around each stimulation frequency.

Signal to Noise Ratio Peak: Signal to Noise Ratio (SNR) is a measure that depends on the frequency f and is computed as the ratio of Fourier Power at frequency f and average Fourier power at its adjacent frequencies. This is actually a way to enhance SSVEP peaks (Wang et al. 2006) and is computed according to the following formula:

$$X'(f) = \frac{nX(f)}{\sum_{k=s}^{n/2} X(f+k\Delta f) + \sum_{k=s}^{n/2} X(f-k\Delta f)}, \quad (1)$$

where X(f) is the value for Fourier power of a signal at the frequency f and X’(f) is the value of the SNR at frequency f, and Δf is the frequency step. The maximum SNR value at a small margin around each stimulation frequency is then defined as the SNR Peak.

We computed Fourier Peak and the SNR Peak for occipital, parieto-occipital and Frontal channels. (O_1, O_2, Po_3, Po_4, F_3, F_4)

Magnitude Squared Coherence: Magnitude Squared Coherence (MSC) is a measure for quantifying the synchronization between two signals. This feature is computed between pairs of EEG channels to see how similar their power spectrums in terms of magnitude are. The magnitude squared coherence is a function of the power spectral densities (Pxx(f) and Pyy(f))and the cross power spectral density (Pxy(f)) of x and y.

$$\mathbf{C_{xy}}(f) = \frac{|\mathbf{P_{xy}}(f)|^2}{\mathbf{P_{xx}}(f)\mathbf{P_{yy}}(f)}, \quad (2)$$

We computed this value for the following channel pairs: O_1- O_2, F_3-F_4, Fp_1-O_1, Fp_2-O_2, Fp_1-O_2, Fp_2-O_1.

Fourier and SNR Peak for concatenated Signals: The FFT epochs of SSVEP signal require sufficient data length to achieve a satisfactory frequency resolution. However, increased epoch length comes at the cost of time taken to collect EEG. Since for the purpose of online BCI applications, time for processing the data and estimation of the command is a crucial element to be kept short, detection should be done using short epochs of signals. Tomita et al. (2011) proposed concatenation method to improve the frequency resolution of the SSVEPs using short time window epochs. In their proposed method, they concatenated signals from different channels in the time domain and showed that the concatenated signal produces clearer SSVEP peaks in the Fourier Spectrum due to the increased frequency resolution.

For this study, two groups of concatenated signals were built, one including two frontal channels (F_3 and F_4) and the other one including the parieto-occipital and occipital channels (O_1, O_2, Po_3 and Po_4). Then the Fourier and SNR Peak were computed for both concatenated signals.

2.4 Feature Ranking

Since the number of candidate features was too large (Nf = 22) given the number of examples in the database (Ne = 100) for each stimulation frequency, feature selection was performed by the orthogonal forward regression (OFR) algorithm (Guyon and Elisseef, 2003) to select the most relevant features for discriminating the 9 Stimuli. For this purpose we used OFR algorithm 36 times, each time finding the most relevant features for discrimination of two different Stimuli (36 different combinations for 9 different stimuli frequencies). OFR algorithm calculates the angle between each candidate feature $\mathbf{u_i}$ and the quantity to be modelled **y** and defines the most correlated feature as the feature that has the smallest angle (Θi) with **y.**

$$\mathbf{u_j} = arg\ (max_i\{cos^2(\Theta_i)\}), \quad (3)$$

Then **y** and all the remaining candidate features are projected onto the null space of the selected feature and the same procedure is iterated until all candidate features are ranked. We performed our feature ranking on half of our database in order to avoid a bias. Finally 8 features were selected (see Table).

Table 1: Top-ranked features from the feature selection step.

Feature	Channels
Frequency peak	Average (O_1,O_2)
Frequency peak	Average (F_3,F_4)
SNR peak	Average (Po_3,Po_4)
Magnitude squared coherence	O_1, O_2
Concatenation: frequency peak	F_3, F_4
Concatenation: SNR peak	F_3, F_4
Concatenation: frequency peak	O_1, O_2,Po_3, Po_4
Concatenation: SNR peak	O_1, O_2,Po_3, Po_4

3 CALIBRATION

3.1 Distribution Calibration

We hope to reduce the inter-subject variability by calibrating the data, so that the SSVEP responses would be more homogeneous. 'First level' calibration is based on a mathematical projection of the data into a reference space, which is defined based on a short period of time. We investigated two different first level calibration approaches:

- Resting state eyes open data. In this case, we remove for each feature the mean value of 30 sec of resting state.
- Resting state eyes open data and active state data (see Figure 2): active state data is a collection of 30 sec of SSVEP response at a given frequency. In this case, we remove the separating threshold between active and passive data, so that active data and non-active data will be discriminated on their sign. The threshold is determined using linear classification (active *vs.* rest data). Active data values should then be positive and non-active data values should be negative.

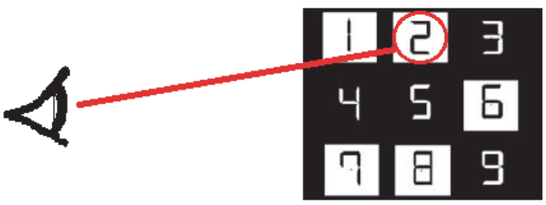

Figure 2: Frequency #2 is defined as active state frequency while the other pads are defined as non-active.

3.2 Feature Calibration

3.2.1 Harmonic Selection

Depending on the subject and the observed frequency of stimulation *f*, features can have higher value depending on if they are calculated at the fundamental *f* or at the harmonic 2*f*. While the mean value is commonly used, a calibration can be performed to detect the frequency and feature that emphasize this specificity for each subject. Based on a 9 30-second recordings for each stimulation frequency, a selection between the value of the feature at f, 2f and the mean value is processed via the Mann-Whitney test. This test determines whether the medians are significantly different. A case where the Mann-Whitney z-score is above 2 or -2 indicates that either the fundamental or the harmonic dominates. Otherwise the medians are not different enough, and then the mean value is kept.

3.2.2 Channel Topography Selection

Eight classic features were isolated as explained above, but these 8 features are not optimal for all the subjects. We optimized the channel selection, by subdividing the selected features into 7 groups of features. The 7 groups are organized in order to access a subject-specific topography (see Table).

Table 2: 7 Groups of expanded features These features are more detailed topographic mappings of the selected features, so that the topography can be further adapted to each subject.

Group1	Concatenation: frequency peak	O_1 O_2 Po_3 O_1 O_2 Po_4 O_1 Po_1 Po_4 O_2 Po_1 Po_4 O_1 O_2 Po_3 Po_4
Group2	Concatenation: SNR peak	O_1 O_2 Po_3 O_1 O_2 Po_4 O_1 Po_3 Po_4 O_2 Po_3 Po_4 O_1 O_2 Po_3 Po_4
Group3	Magnitude squared coherence 1	O_1 O_2 Po_3 Po_4 O_2 Po_3 O_1 Po_4
Group4	SNR peak	F_3 F_4 F_3 F_4
Group5	SNR peak	O_1 O_2 Po_3 Po_4 Po_3 O_2 Po_4 O_1 O_1 O_2 Po_4 Po_3 O_1 O_2 Po_3 Po_4
Group6	Magnitude squared coherence 2	F_3 O_2 F_3 Po_4 F_4 O_1 F_4 Po_3
Group7	Frequency peak	O_1 O_2 O_1 O_2

Based on one 30-second recording at only one

stimulation frequency, considered as a reference, one feature of each group is selected via OFR for each subject (therefore personalizing the channel topography for each subject). Each subject has then his personalized7-feature set.
The selection processed is OFR based:

- find the best feature in the 33 features.
- remove all features corresponding to the same group
- project the remaining candidate features (from the other groups) onto the null space of the selected feature.

The above two steps can be iterated in subspaces of decreasing dimensions until one candidate of each group has been selected.

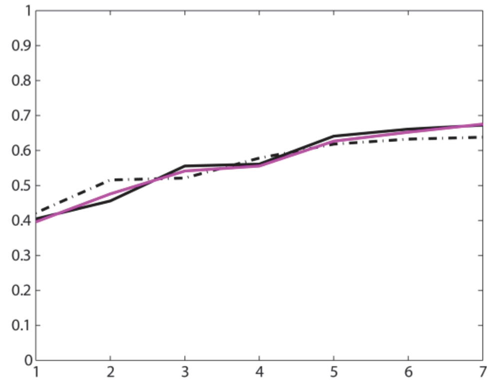

Figure 3: SR rate, sorted from the worst to the best subject. Black: no calibration (mean = 0.56); black dotted: passive distribution calibration (mean = 0.56); magenta: active distribution calibration (reference frequency 6.67Hz, mean = 0.56).

4 CLASSIFICATION

The classes correspond to the frequencies of stimulation, they indicate which of the buttons on the dial pad the subject wishes to activate. The 9-class classifier is in fact composed of 36 2-class linear classifiers (LDA). Each 2-class classifier is based on twice the features account, as the class parameters are the features calculated at both frequencies of analysis.

For a given 3s signal, the features are extracted for the 9 stimulation frequencies. Then for each couple (f_a, f_b) of frequencies, the corresponding features are compared to determine whether the signal corresponds to class #a or class #b, in other words if the observed command flickered at frequency f_a or at f_b. The command estimation is based on the best mean score after the 36 comparisons.

5 RESULTS

5.1 Distribution Calibration

We classify the data using a cross validation approach. As we intend to test the capability of the system to adapt to new subjects, we iteratively remove one subject, train the classifier with the other data, and test on the rejected subject. This method is similar, in spirit, with the classical leave-one-out cross-validation approach – but here we leave one subject out, instead of only one example.

The success rate (SR) is defined as the trace of the confusion matrix after Leave-One-Subject-Out testing. Generally, SR is not significantly improved by the distribution calibration, except for the calibration on rest data which slightly improved the worst subject (at the expense of decreasing the SR for the best subject). These classification results are stable across frequencies, as is illustrated on the confusion matrix of Figure 4.

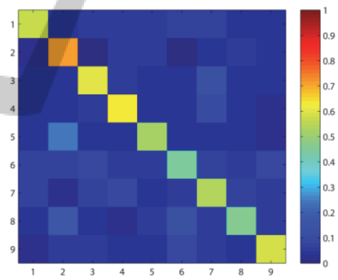

Figure 4: Confusion Matrix (average of subjects). horizontal axis: stimulation frequency. Vertical axis: estimated command.

This confusion matrix corresponds to classification of the data without calibration; the (not shown) confusion matrices for all type of calibrations investigated in this manuscript share the same stability properties across frequencies.

5.2 Feature Calibration

Calibration based on the selection of the dominant harmonic led to a general improvement of the SR, where six of the seven subjects had SR above 0.55 (Figure).

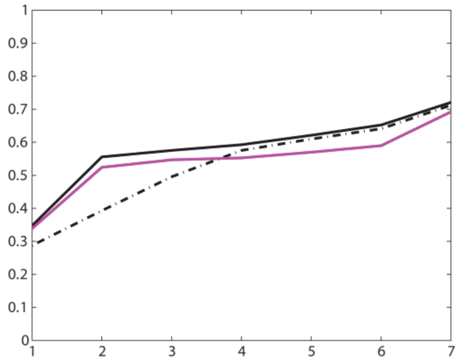

Figure 5: SR rate, sorted from the worst to the best subject. black: harmonic selection (mean = 0.58); black dotted: harmonic selection combined with passive distribution calibration (mean = 0.53); magenta: harmonic selection combined with active distribution calibration (reference frequency 6.67Hz, mean = 0.54).

Topographic selection calibration also led to a significant improvement (average SR = 0.58 for reference frequency 8.75 Hz), but this improvement was not stable across frequencies.

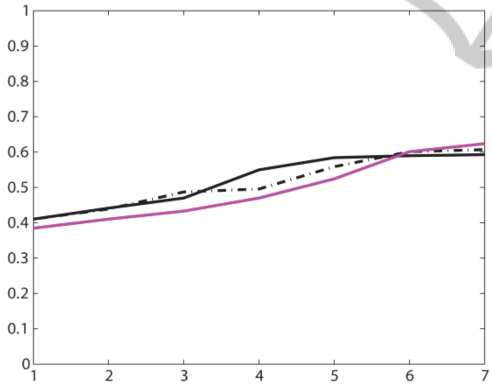

Figure 6: SR rate, sorted from the worst to the best subject. black: topographic selection, using a 4 Hz reference signal (mean = 0.60); black dotted: topographic selection combined with passive distribution calibration (mean = 0.57); magenta: topographic selection combined with active distribution calibration (mean = 0.59).

Combining harmonic selection and topographic selection led to an improvement, which was this time much more stable across reference frequencies (SR = 0.57 for 10 Hz, SR = 0.58 for 6.67 Hz, and SR = 0.60 for 4 Hz).

Whether using harmonic selection, topographic selection, or both, subsequent distribution calibration did not provide any improvement.

6 DISCUSSION

We investigated four different types of BCI system calibration, based on:

- Distribution mapping, using a rest condition signal as reference,
- Distribution mapping, using a rest condition signal and an active signal as references,
- Subject dependent choice of electrodes,
- Subject dependent choice of SSVEP harmonics.

We compared the classification results for the detection of SSVEP peaks of these four calibration methods.

For the choice of stimulation frequencies, we embedded 20 Hz stimulation in the design of our stimulation interface. The SSVEP responses that were generated with this frequency were strong enough to be detected. However, in a study by Bakardjian et al. (2010), the best choice of stimulation frequency for evoking the strongest response is reported to be among 5.6 to 15.3 Hz. On the other hand, there exist studies supporting the usefulness of high frequency stimuli in generating good SSVEP responses. (Wang et al., 2006, Volosyak et al., 2010). Wang et al. (2006) also employed high frequency stimuli in their experimental design, and explained this increase in the stimulus frequency bandwidth not only as a factor to decrease time length of signal epochs for detection of SSVEPs but also as a factor for reducing the eyestrain effect caused by the flickers. However, these effects may vary from subject-to-subject, and we did not investigate further in this direction.

We consistently show that distribution mapping proved to be useless. It never improved the classification rates, whether used alone or in combination with the other two calibration methods. The features values of resting state data don't seem to be comparable to the features value of non-active state frequency. Depending on the subject, the frequency and the features, they can be lower or higher. Thus no way was found yet to find a generalization rule for all subjects. This result could be due to several reasons. First of all, 30 seconds of data may be insufficient to extract a sufficiently stable signature of the EEG activity. Using longer epochs might provide better results. Second, owing to the non-stationary nature of EEG (see *e.g.* Kaplan et al., 2005), it might be necessary to monitor the signal evolution along time (the data collection lasted up to one hour), otherwise the reference data used for calibration may not be a good reference anymore.

Selection of harmonics and topography led to much more clear improvements. This is to be expected: this method seeks to adapt the system to the specificities of each subject. It is noteworthy that each subject has specific brain responses to SSVEP (see *e.g.* Silberstein, et al., 1990), whether topographically or frequency-wise. It is therefore not surprising that an adaptation of the system to the specificities of each subject leads to an improved classification. The best calibration method between those two, according to our results, is the selection of the dominant harmonic in the SSVEP response. However, the reader should keep in mind that those two methods are based on very different approaches. Harmonic selection used 5 minutes of data, whereas topography selection used only 1 minute of data, but still led to some significant improvements. Our results therefore also confirm the interest of selecting the channels, which was already pointed by Wang et al. (2006).

ACKNOWLEDGEMENTS

In this project Hovagim Bakardjian was supported by an International Neuroinformatics Coordinating Facility grant (june 2011), and by the #3862 fellowship of the Fondation Pierre Gilles de Gennes.

REFERENCES

Bakardjian, H., Tanaka, T., Cichocki, A. 2010, Optimization of SSVEP brain responses with application to eight-command Brain–Computer Interface. *Neuroscience Letters*, 469(1):34-38.

Bin, G., Gao, X., Yan, Z., Hong, B., Gao, S., 2009. An online multi-channel SSVEP based brain–computer interface using a canonical correlation analysis method. *J. Neural Eng.* 6 (4) 6: 046002

Friman, O., Volosyak, I., Graser, A., 2007. Multiple channel detection of steady-state visual evoked potentials for brain–computer interfaces. *IEEE Trans. Biomed. Eng.* 54 (4), 742–750.

Kaplan A. Y., Fingelkurts A. A., Fingelkurts A. A., Borisov S. V., Darkhovsky B. S., 2005, Nonstationary nature of the brain activity as revealed by EEG/MEG: Methodological, practical and conceptual challenges. *Signal Processing*, 85(11):2190-2212.

Krauledat, M., Tangermann, M., Blankertz, B., Muller, K. R., 2008. Towards zero training for brain–computer interfacing. *PLoS ONE* 3(8), e2967.

Lin, Z., Zhang, C., Wu, W., Gao, X., 2006. Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs. *IEEE Trans. Biomed Eng.* 532610-4.

Lotte, F., 2011. Generating Artificial EEG Signals To Reduce BCI Calibration Time*, 5th International Brain-Computer Interface Workshop*

Nicolelis M., 2011, Beyond Boundaries: The New Neuroscience of Connecting Brains With Machines-And How It Will Change Our Lives. *Times books*.

Parini, S., Maggi, L., Turconi, A. C., and Andreoni, G., 2009. A Robust and SelfPaced BCI System Based on a Four Class SSVEP Paradigm: Algorithms and Protocols for a High-Transfer-Rate Direct Brain Communication*. Comput.Intelli. Neurosci* ID 864564.

Silberstein, R. B., Schier, M. A., Pipingas, A., Ciorciari, J., Wood, S. R., Simpson, D. G., 1990. Steady-state Visually evoked potential topography associated with a visual vigilance task*. Brain Topogr*. 3, 337–347.

Shishkin S. L., Nikolaev A. A., Nuzhdin Y. O., Zhigalov A. Y., Ganin I. P., Kaplan A., Y. 2011. Calibration of the P300 BCI with the single-stimulus protocol. *Proc. of the Fifth International BCI Conference Graz University of Technology*, Austria. 256-259

Tomita, Y., Gaume, A., Bakardijian, H., Maurice, M., Cichocki, A., Yamaguchi, Y., Drefus, G., Vialatte, F., 2011. Concatenation Method for high temporal resolution SSVEP-BCI. *International Conference on Neural Computation Theory and Applications*.

Vialatte, F. B., Maurice, M., Dauwels, J., Cichocki, A., 2010. Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives, *Progress in Neurobiology* 90 418–438

Volosyak, I., Malechka, T., Valbuenaand, D., Graser, A., 2010. A Novel Calibration Method for SSVEPBased Brain Computer Interfaces. *18th European Signal Processing Conference*, 939-943.

Wang, Y., Wang, R., Gao, X., Hong, B., Gao, S., 2006. A practical VEP-based brain–computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng*. 14 (2), 234–239.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., Vaughan, T. M., 2002. Brain–computer interfaces for communication and control. *Clin. Neurophysiol*. 113 (6), 767–791.

Zhang, Y., Jin, J., Qing, X., Wang, B., Wang, X., 2012. LASSO based stimulus frequency recognition model for SSVEP BCIs, *Biomedical Signal Processing and Control* (7)104– 111