Towards an SSVEP Based BCI With High ITR

Ivan Volosyak, Diana Valbuena, Thorsten Lüth, and Axel Gräser

Abstract

A brain-computer interface (BCI) provides the possibility to translate brain neural activity patterns into control commands without movement by the user. In recent years, there has been increasing interest in using steady-state visual evoked potential (SSVEP) in BCI systems; the SSVEP approach provides currently the fastest and most reliable communication paradigm for the implementation of a non-invasive BCI system. However, many aspects of current system realizations need improvement, specifically in relation to speed (in terms of information transfer rate as well as time needed to perform a single command), user variability and ease of use. With these improvements in mind, this paper presents the Bremen-BCI, an online multi-channel SSVEP-based BCI system that operates on a conventional computer making use of the minimum energy combination method for extraction of power information associated with the SSVEP responses. An additional advantage of the presented methodology is that it is fully online, i.e., no calibration data for noise estimation, feature extraction, or electrode selection is needed, the system is ready to use once the subject is prepared. The SSVEP-based Bremen-BCI system with five targets, an adaptive time segment length between 0.75s and 4s, and six EEG channel locations on the occipital area, was used for online testing on 27 subjects. ALL participants were able to successfully complete spelling tasks with a mean accuracy of 93.83% and an information transfer rate (ITR) of 49.93 bit/min.

Index Terms

BCI - Brain-Computer Interface, EEG - Electroencephalogram, SSVEP - Steady-State Visual Evoked Potential, ITR - Information Transfer Rate, LCD - Liquid Crystal Display.

I. Introduction

Brain-computer interface (BCI) systems allow people to interact with the environment through an alternative communication channel entirely independent from the traditional motor output pathways of the nervous system [1]. These devices may be the only possible way of communication for severely disabled users, such as persons suffering from Cerebral Palsy, stroke victims, or persons with injuries of the brain or spinal cord. Recent studies have indicated an increased interest in non-invasive BCI systems which are based on various sensory modalities [2]. In non-invasive BCIs, electroencephalography (EEG) is commonly used because of its high time resolution, ease of acquisition, and lower system cost as compared to other brain activity monitoring modalities.

Manuscript received October 18, 2010. The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant BRAIN, n° 224156, from a Marie Curie European Re-Integration Grant RehaBCI, n° 224753, and by the German Federation of Industrial Research Associations (AiF) under Grant sBCI, 16136BG.

The authors are with the Institute of Automation, University of Bremen, 28359 Bremen, Germany. Corresponding author is Axel Gräser, ag@iat.uni-bremen.de

Nowadays, the most commonly used EEG-based BCI systems employ event-related synchronization of mu and beta bands, event-related potentials and steady-state visual evoked potential (SSVEP). The performance of the BCI can be assessed by the information transfer rate (ITR) as introduced in [3] and reported in the majority of BCI studies. This measure depends on three factors: speed, accuracy and number of targets, which can vary from two [4] up to 48 [5]. The SSVEP approach provides currently the fastest and the most reliable communication paradigm for the implementation of a non-invasive BCI system [6], [7]. In a six target SSVEP-based BCI, an average accuracy of 95.3% and information transfer rates of 58 ± 9.6 bit/min for 12 healthy participants were reported in [8]. Other studies have reported classification accuracies of more than 90% [9], [10].

High information transfer rates are essential for a BCI in order to become a practical device. Because of the very low information throughput of existing BCI systems, the majority of research groups have paid special attention mostly on improvements of data acquisition techniques and signal processing methods, while the human factor in the human-machine interaction has been typically neglected. This aspect will become more and more important as the ITRs increase, as the user must be able to cope with the high speed of the overall system. Already ITR values of more than 30 bits/min can cause problems for the user in the free spelling mode where the user is requested to decide what to spell first, which is typically associated with a high cognitive load for the user.

Another key aspect of a practical BCI system is a friendly graphical user interface (GUI), which should be simple and intuitive. The use of the LCD screen on conventional PCs for creating the visual stimuli for SSVEP based BCIs will lower the system cost and allow the development of portable BCI systems. Besides that, the presentation of stimuli directly on the LCD screen allows displaying the current state of the control signals in the form of real-time visual feedback.

This work is one of the first examples of a practical BCI. Our goal was to build a BCI with high ITR to provide a faster communication between man and machine. The approach was the improvement of two important aspects of the BCI system, the signal processing and the real-time visual feedback presented to the user, in order to achieve higher information transfer rate than the currently available. These improvements could be used for the practical design of other kinds of BCI systems, not only for the systems based on the SSVEP paradigm.

The paper is organized as follows. The second section discusses the details of the proposed improvements in practically all parts of the SSVEP-based BCI system. The achieved results are presented in the third section, followed by discussion and conclusion in two final sections.

II. Materials and Methods

A. SSVEP Signal Processing

The SSVEP signal detection and classification methods are the core of this work. The minimum energy combination (MEC) method [11] was used to create a spatial filter that magnifies the SSVEP response and cancels nuisance signals and noise. The BCI automatically determined the best spatial filter for each subject at each stimulation frequency. SSVEP detection is based on power estimation after spatial filtering and a statistical probability method that enhances signal separability. Moreover, an adaptive mechanism is used for selection of the appropriate window length that depends on the subject's online performance. This classification algorithms were implemented in C++ building an asynchronous, real-time BCI system. The complete signal processing approach used for online classification of SSVEP responses is summarized as follows.

1) SSVEP Response and Modeling: An SSVEP-BCI reflects the user's attention to an oscillating visual stimulus. The stimuli which are commonly used are light sources flickering at different frequencies. They elicit responses in the visual cortex of the brain, corresponding to SSVEPs at the same frequencies and their higher harmonics. The amplitude and the phase that define an SSVEP response depend on the frequency, intensity and the structure of the repetitive visual pattern [12]. It is possible to obtain an SSVEP response at a large range of frequencies, from 1 Hz to 90 Hz [13], however the strongest responses are typically obtained for lower stimulation frequencies around 15 Hz [14].

To model an SSVEP response, we consider a visual stimulation with a flicker-frequency of f Hz. The voltage between the i-th electrode and a reference electrode at time t, $y_i(t)$, can then be described as a function of the stimulus frequency, f, and its harmonics, subject to a phase-shift, and a noise and nuisance signal, $E_{i,t}$,

$$y_i(t) = \sum_{k=1}^{N_h} (a_{i,k} \, \sin 2\pi kft + b_{i,k} \, \cos 2\pi kft) + E_{i,t} \tag{1}$$

where N_h is the number of considered harmonics. The model is linear and the signal is composed of two parts. The first part corresponds to the visually evoked response signal, which is composed of a number of sine and cosine functions at the harmonic frequencies kf with specific amplitudes, $a_{i,k}$ and $b_{i,k}$. The second part of the model, $E_{i,t}$, represents all the information that cannot be attributed to the SSVEP response such as environmental noise and its effect on the subject, and natural physical disturbances like other background brain processes and various kinds of artifacts.

For a time segment of length T_s, acquired with a sampling frequency of F_s Hz, which contains N_t samples of the ith signal, the model can be expressed in vector form as follows

$$\boldsymbol{y}_i = \boldsymbol{X}\boldsymbol{g}_i + \boldsymbol{E}_i \tag{2}$$

where $\boldsymbol{y}_i = [y_i(1), \ldots, y_i(N_t)]^T$ contains the EEG signal for the electrode i in the time segment used for the signal analysis. The SSVEP information matrix $\boldsymbol{X}$ is of size $N_t \times 2N_h$ and contains the sine and cosine components associated with the N_h harmonics, while the vector $\boldsymbol{g}_i$ of size $2N_h \times 1$ contains the corresponding amplitudes $a_{i,k}$ and $b_{i,k}$. Equation (2) can be generalized for N_y electrodes,

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{G} + \boldsymbol{E} \tag{3}$$

where $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{N_y}]$ contains the sampled EEG signals from all the electrodes. The matrix $\boldsymbol{G}$ of size $2N_h \times N_y$ contains all the amplitudes for all the expected sinusoids for all electrode signals.

2) Minimum Energy Combination: To extract discriminant features, the signals from the i electrodes need to be combined. This can be achieved by defining a channel vector $\boldsymbol{s}$ of length N_t which is a linear combination of the electrode signals, $\boldsymbol{y}_i$,

$$\boldsymbol{s} = \sum_{i=1}^{N_y} w_i \boldsymbol{y}_i = \boldsymbol{Y}\boldsymbol{w} \tag{4}$$

where $\boldsymbol{w}$ is a vector of weights $[w_1, \ldots, w_{N_y}]$ associated with the individual electrode signals. The aim of the channel $\boldsymbol{s}$ is to enhance the information contained in the EEG while reducing the nuisance signals. Several channels can be created by using different sets of weights, depending on the nature of the SSVEP signal and the noise. Equation (4) can be generalized for N_s channels as

$$\boldsymbol{S} = \boldsymbol{Y}\boldsymbol{W} \tag{5}$$

with the set of channels $\boldsymbol{S} = [\boldsymbol{s}_1, \ldots, \boldsymbol{s}_{N_s}]$ and the corresponding weight matrix $\boldsymbol{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{Ns}]$.

As a first step, an orthogonal projection is used to remove any potential SSVEP activity from the recorded signal,

$$\tilde{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y} \tag{6}$$

The remaining signal $\tilde{\boldsymbol{Y}}$ contains approximately only noise, artifacts and background activity.

In the next step the weight vector $\hat{\boldsymbol{w}}$ is found which minimizes the energy of the signal $\tilde{\boldsymbol{Y}}$, by optimizing

$$\min_{\hat{\boldsymbol{w}}} \left\| \tilde{\boldsymbol{Y}}\hat{\boldsymbol{w}} \right\|^2 = \min_{\hat{\boldsymbol{w}}} \hat{\boldsymbol{w}}^T \tilde{\boldsymbol{Y}}^T \tilde{\boldsymbol{Y}} \hat{\boldsymbol{w}} \tag{7}$$

which will minimise the component of the noise and nuisance signal in the corresponding channel signal (equation (4)). As shown in [11], the lower bound of the quadratic form on the right hand side in equation (7) is given by the minimal eigenvalue λ_1 of the matrix $\tilde{\boldsymbol{Y}}^T\tilde{\boldsymbol{Y}}$. The solution is therefore the corresponding eigenvector, $\boldsymbol{v}_1$, which gives the weight vector for one channel. Additional uncorrelated channels can be added by choosing the next smallest eigenvalue (and corresponding eigenvector). The weight matrix can therefore be chosen based on the eigenvalues in ascending order $(\lambda_1, \lambda_2, \ldots)$ and the corresponding eigenvectors $(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots)$,

$$\boldsymbol{W} = \left[\frac{\boldsymbol{v}_1}{\sqrt{\lambda_1}} \;\cdots\; \frac{\boldsymbol{v}_{N_s}}{\sqrt{\lambda_{N_s}}} \right] \tag{8}$$

The total number of channels used, N_s, is selected by finding the smallest value for N_s which satisfies

$$\frac{\sum_{i=1}^{N_s} \lambda_i}{\sum_{j=1}^{N_y} \lambda_j} > 0.1 \tag{9}$$

This can be interpreted as selecting the number of channels in such a way as to discard as close to 90% of the nuisance signal energy as possible [11].

3) SSVEP Detection: To detect the presence of a frequency in the spatially filtered signals, the power in that frequency and a number of harmonics N_h ($N_h = 2$ in the actual system implementation to avoid overlapping between the frequencies used) can be estimated by,

$$\hat{\boldsymbol{P}} = \frac{1}{N_s N_h} \sum_{l=1}^{N_s} \sum_{k=1}^{N_h} \left\| X_k^T s_l \right\|^2 \tag{10}$$

In a BCI application, only the frequencies with which the user is stimulated should produced a control signal. To improve the robustness of the classification, not only the stimulation frequencies are detected, but a number of additional frequencies. As the quality of the SSVEP response depends on the stability of the frequencies, the five stimulation frequencies that are used in this experiment were selected based on the refresh rate of the LCD screen (120 Hz) that produces the stimuli: 6.67 Hz ("select"), 7.50 Hz ("left"), 8.57 Hz ("right"), 10.00 Hz ("up"), and 12.00 Hz ("down"). These frequencies correspond to periods equivalent to a fixed number of frames on the LCD screen that assures frequency stability [15], [16]. We consider four additional frequencies to improve the reliability of the outputs: 7.08, 8.03, 9.28, and 11.00 Hz. These frequencies are selected between two target frequencies. For instance, 7.08 Hz is the mean value of 6.66 and 7.50. The purpose of these frequencies is to improve the quality value of the detection of the frequencies of interest as described below.

The SSVEP power estimations for all frequencies N_f, in the considered case $N_f = 9$, are normalized into powers:

$$p_i = \frac{\hat{P}_i}{\sum_{j=1}^{j=N_f} \hat{P}_j} \quad \text{with} \quad \sum_{i=1}^{i=N_f} p_i = 1 \tag{11}$$

where $\hat{P}_i$ is the *i*-th signal power estimation, $1 \leq i \leq N_f$.

A high normalized power will become more difficult to achieve when N_f is large (i.e., adding other frequencies will amplify this effect). Besides, we use a Softmax function to enhance the gap between the values:

$$p'_i = \frac{e^{\alpha p_i}}{\sum_{j=1}^{j=N_f} e^{\alpha p_j}} \quad \text{with} \quad \sum_{i=1}^{i=N_f} p'_i = 1 \tag{12}$$

where α is set to 0.25 (based on our prior practical investigations and on the number of frequencies used N_f).

Although this function does not change the distribution of the frequency powers, it improves the relevance of the command detection.

4) Signal Classification: The classifier output O is determined as the number of the *i*-th frequency, if 1) this *i*-th frequency has the highest probability p'_i, 2) p'_i exceed the pre-defined threshold β, and 3) the detected frequency belongs to one of the stimulating frequencies:

$$O = \begin{cases} argmax_i(p'_i), \\ p'_i \geq \beta, \\ i \leq 5 \end{cases} \tag{13}$$

where $1 \leq i \leq N_f$ and β is set to 0.35. This choice of β is based on our prior practical investigations and on the number of used frequencies N_f.

If O is classified as an undesired frequency ($i > 5$) then this classification will be rejected as the detected frequency does not belong to the expected frequency set. To improve the overall reliability of the system, the commands corresponding to the stimulating frequencies are produced only if their probability is higher than the fixed threshold β.

The main advantage of the methodology outlined above is the fact that the pre-defined threshold β represents the relative value and not the absolute value as in the original method [11], [17], and as such it is independent of changes in the segment length T_s of the acquired EEG signal used for classification.

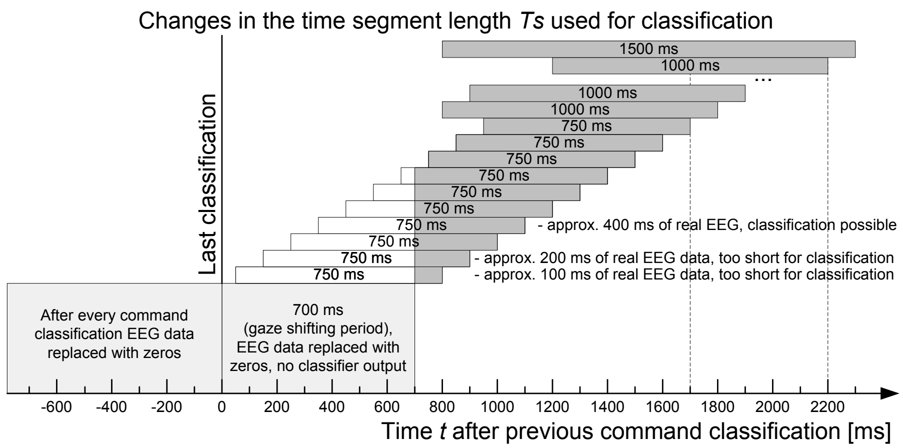


Fig. 1. Changes in the time segment length (T_{s_n}) after each performed classification in case where no classification can be made at the moment and the actual time t allows the extension of the T_s to the next pre-defined value (considering the gaze shifting period of 700 ms). The horizontal bars represent the EEG data used for classification at the time t after the previous command classification (gray areas represent the real EEG data, the "zero out" EEG is shown in white).

The classification has to take into account the moment when the subject does not focus on any stimuli. Therefore, the classifier output detects a resting state or transition states between two stimuli and moments when the user's attention is not on any particular stimulus. These states where no SSVEP response should be detected are referred to as zero class. If stimulation frequencies are located at the alpha band, this could produce false classifications in resting state.

5) Adaptive Mechanism: In addition to the already applied online adaptation of the MEC, in which the number of channels used is recalculated every 100 ms, the possibility of online adaptation of the time segment length T_s used for classification allows further improvements. At the beginning of the trial, the first classification will be performed with the minimal time segment length of 750 ms (the values of 750, 1000, 1500, 2000, 3000 and 4000 ms were determined during our previous work, for more details please refer to [18]). Similar to the original method, the classification is performed with the sliding window of T_s based on the last acquired EEG data approx. every 100 ms (every 13 samples with the sampling rate of 128 Hz used). In the case where no classification can be made and the actual time t allows the extension of the T_s to the next pre-defined value, this new value will be used instead:

$$\forall t : T_s = \begin{cases} 750\text{ms}, & t \leq 1000\text{ms} \\ 1000\text{ms}, & 1000\text{ms} < t \leq 1500\text{ms} \\ \dots & \\ T_{s_n}, & T_{s_n} < t \leq T_{s_{n+1}} \\ \dots & \\ 4000\text{ms}, & t > 4000\text{ms} \end{cases} \tag{14}$$

Further, after each performed classification, the EEG data used for the classification (T_{s_n}) are empty (replacement with zeros). This ensures that no subsequent classification of the same frequency based on the buffered data can be done. In addition, the user of the BCI system will need some time for the gaze shifting. An additional time of 700 ms of the EEG data after each classification will not be utilized, since they are usually contaminated by strong movement artifacts. Therefore this data exclusion (replacement with zeros, reset the classifier output) is helpful for the reliable classification. The next classification will be performed (after the end of gaze shifting period, $t = 0$ in eq. (14)) with the minimal value T_s=750 ms. Equation (14) could be represented as follows (considering the time shift of 700 ms):

$$\forall t : T_s = \begin{cases} 750\text{ms}, & t \leq 1700\text{ms} \\ 1000\text{ms}, & 1700\text{ms} < t \leq 2200\text{ms} \\ \dots & \\ T_{s_n}, & T_{s_n} < t \leq T_{s_{n+1}} \\ \dots & \\ 4000\text{ms}, & t > 4700\text{ms} \end{cases} \tag{15}$$

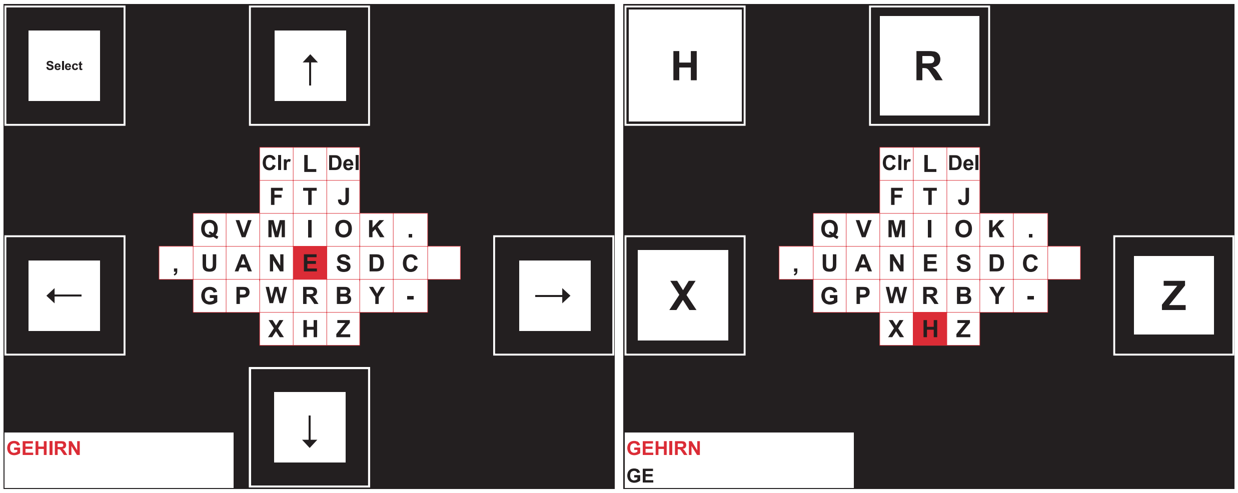
Fig. 1 shows the change in the used time segment length after performed classification. The next classification could be performed again approx. 400 ms after the end of gaze shifting period (subject dependent). That means 400 ms + 700 ms after the previous command classification, as during this time the EEG data are only partially available (for this example, a time segment length of 750 ms is the sum of 400 ms of real EEG data and 350 ms of zeros from gaze shifting period). In practice, the minimal time between two commands classifications was approx. 1050 ms.

Please note, in order to provide the overall stability, there is no changes in the overall synchronization in the BCI system - the classifications are performed all the time approx. every 100 ms (every 13 samples).

B. Bremen-BCI spelling application

The graphical user interface (GUI) of the Bremen-BCI speller is presented in Fig. 2. It consists of a virtual keyboard with 32 characters (letters and special symbols), and five white stimulation boxes. These boxes are at the outer edges and upper left corner of the screen and flicker with different frequencies that correspond to the commands “left”, “right”, “up”, “down”, and “select”. This setup, as opposed to having an LCD for the GUI and a separate LED board for the visual stimuli, is much more convenient for the user as they do not have to shift their gaze too much. Further details about the software design and implementation of this GUI application can be found in [17], [19].

At the beginning of each trial, the cursor is located in the middle of the virtual keyboard over the letter ‘E’ and all oscillating boxes are presented in their default size (150 x 150 pixels), as shown in Fig. 2a. During the spelling task, by focusing on one of the four oscillating boxes, the cursor is navigated by the commands “left”, “right”, “up” and “down” until the desired letter is reached. With the “select” command a letter is selected and displayed at the bottom of the screen, as shown in Fig. 2b. Audio feedback follows every recognized command. After every selection the cursor automatically moves back to the initial letter ‘E’. A minimum of 9 commands is needed in



(a) At the beginning of the experiment, all flickering boxes are represented in their default size of 150x150 pixels. The arrows on the flickering boxes and the command "Select" are self-explaining (intuitive GUI).

(b) During the experiment, subject is spelling the word "GEHIRN." The actual letter "H" is going to be selected. As no further movement in the "down" direction is possible, this stimulus has been deactivated.

Fig. 2. GUI of the SSVEP based BREMEN-BCI on a LCD screen (22" Samsung SyncMaster 2233 with the vertical refresh rate of 120 Hz and resolution of 1680 x 1050 pixels). A cursor can be navigated left, right, up and down until the desired letter is reached. With the "select" command a letter is selected and displayed at the bottom of the screen. After every selection, the cursor automatically moves back to the initial letter 'E'. Therefore, a minimum of 9 commands is needed in order to spell the word "BCI" ("down", "right", "select", "right", "right", "right", "select", "up", "select").

order to spell the word "BCI." The letter 'B' can therefore be reached in two different ways: 1) "down", "right" or 2) "right", "down". Both paths are judged as correct command sequences.

During the experiment, oscillating boxes can be enlarged in relation to the SSVEP amplitude. The white frames around each stimulus box (fixed sizes of 250 x 250 pixels) represents the maximum size that a stimuli can reach. This helps the user to know whether a command is executed. This novel continuous real-time visual feedback about the power of SSVEP signals improves further the time behavior of the BCI system. In general, the amplitude of SSVEP response depends on the size of the visual stimuli, larger stimuli produce better responses. These changes represent the positive feedback in the overall BCI system. Additionally, after successful command execution the next target for each navigation command is displayed on the corresponding flickering box (see Fig. 2b). Navigation cannot move beyond the layout boundaries. As for example, it is not possible to go from the letter 'G' to the letter 'Q' by choosing the "down" command. Since, the corresponding "down" stimulus is hidden. All these modifications improve the comfort and easy of use of the SSVEP-based Bremen-BCI and also increase the overall reliability of the system. The box at the bottom of the screen contains for the copy spelling mode the word to spell and the actual spelled text. Fig. 2b shows a screenshot taken during the online spelling task, when a subject spelled the word "GEHIRN" (German for "brain"). After the subject had successfully spelled the letters "GE", the cursor is navigated over the next character, the letter 'H'. As no further movement in the "down" direction is possible, this stimulus is deactivated.

C. Subjects

A total of 27 subjects participated in the study. Subjects mean age was 23.59 years, range 18-35 with standard deviation 4.73. This study included a total of 21 naive subjects who had never used any kind of BCI system before, subjects 1-3 had extensive SSVEP-BCI experience and were included in this study in order to form the reference value, subject 27 used the Bremen-BCI system once, and subjects 5 and 15 were "BCI illiterates" during the previous experiments. None of the subjects had neurological or visual disorders. Spectacles were worn when appropriate. Subjects did not receive any financial reward for participating in this study.

D. Data acquisition

The experiments were carried out in a normal office room in the Institute of Automation at the University of Bremen. It is worth noting that this is different to the usual EEG recording conditions which is usually an electrically shielded room with low background noise and luminance. Subjects were seated in a comfortable chair approximately 60 cm from a LCD monitor with the graphical user interface shown in Fig. 2. The EEG data were recorded from the surface of the scalp via eight sintered Ag/Ag-Cl EEG electrodes. They are placed on AF_Z for ground, right ear lobe was used for the reference electrode and $P_Z, PO_3, PO_4, O_Z, O_9, O_{10}$ as the input electrodes on the international system of EEG measurement. Standard abrasive electrolytic electrode gel was applied between the electrodes and the skin to bring impedances below $5k\Omega$. An EEG amplifier g.USBamp (Guger Technologies, Graz, Austria) was used for these experiments. The sampling frequency was 128 Hz. During the EEG acquisition, an analog bandpass filter between 2 and 30 Hz, and a notch filter around 50 Hz (mains frequency in Europe) were applied directly in the amplifier.

E. Procedure

Each subject completed a brief questionnaire including the age and gender information and was prepared for EEG recording. Next, a short familiarization run was carried out in order to introduce the experimental procedures and the letters arrangement. The assessment task was to spell five messages with the SSVEP based Bremen-BCI system. Three of the messages were the same for all subjects and were chosen by the experimenter (copy spelling), and two words were chosen by the subject (free spelling). The copy spelling words were "BCI", "GEHIRN" (German for "brain"), and "INTERFACE." The subjects were told that their first free spelling word should be five letters long and their second word does not have any restrictions. Free spelling tasks are referred to further in this paper as "FREE1", which means a word composed of five letters, and "FREE2" for a word without restrictions. Before the free spelling trial, each subject verbally told the experimenter the phrases that she/he intended to spell. The order in which these five phrases were presented to the user was determined randomly to avoid adaptations. Each trial ended automatically, when the subject correctly spelled the specified word (or when the subject chose to stop spelling due to any reason such as visual fatigue - this happened to subjects 15 and 20). Misspellings should be corrected with the 'Del' option located at the top-right of the matrix. At the end of each session, after the phrases were spelled, subjects completed a second questionnaire and the procedure was complete. The entire session took on average about 40 minutes per subject.

F. ITR calculation

The ITR calculation lead to the shown results is based on the formula presented in [3].

$$B_t = \log_2 N + A \log_2 A + (1 - A) \log_2 \left[\frac{1 - A}{N - 1} \right] \tag{16}$$

where A is the classification accuracy and N is the number of targets. B_t is calculated in bits per trial. It is important to note that the number of targets in our case is the number of flickering boxes and not the number of letters and special characters in the letters layout, because none of the letters and special characters are flickering and therefore cannot be directly selected by means of SSVEP communication channel. The number of commands required for the letter selection vary from letter to letter from the minimum one (selection of the letter 'E' in the middle of the speller layout with just one "select" command) to maximum five commands (e.g., selection of the letter 'G', four movement commands and the following selection).

In the case that one wrong moving command is detected, the user should correct this error first, e.g. the correction movement command "right" after the erroneously detected "left" command. Therefore, in this case the correction step is counted as correct command classification. So that, the number of commands can increase depending of the subject's performance. In case of incorrectly classified selection command, the wrongly spelled letter should be corrected, this results in five additional commands to select the special character "Del".

Classification accuracy A is calculated in traditional way and is defined as the number of correct command classifications divided by the total number of classified commands. If some of the stimuli are deactivated, these frequencies still can be classified due to various reasons like e.g. background brain activity in the alpha range (8-12 Hz). Since all five frequencies can be (erroneously) classified independently of the actual cursor position, the assumption that all choices are equally probable still could be suggested.

The spelling time T (for the whole word) is considered in the calculation of the ITR in bits per minute (B_m). Since some of the flickering boxes were deactivated when the current cursor position was on an edge of the speller layout, the ITR calculations considers $N = 2, 3, 4,$ or 5 depending on the cursor position. This leads to the modified ITR calculation:

$$B_m = \frac{60}{T} \cdot \sum_{N=2}^{5} [C_N \cdot B_t(N)] \tag{17}$$

where C_N is the number of classifications at N targets and T is the spelling time in seconds. B_t in this case is a function of N. It is worth noting that this at first sight quite complicated ITR calculation results only in very minor decrease of the ITR values. This could be easily explained with the selected speller layout: only the letters 'G', 'H', 'F' from altogether 18 letters of the three copy spelling words "BCI", "GEHIRN", and "INTERFACE" are located at the layout boundaries, where some stimuli will be deactivated (N = 3 for 'G', N = 4 for 'H', and N = 4 for 'F'). Since during the correct spelling of the word "BCI" the cursor is never located at the layout boundaries, the number of targets N = 5 during the complete spelling task "BCI", C_2 = 0, C_3 = 0, C_4 = 0, C_5 = 9, respectively, and the ITR is calculated in the conventional way:

$$B_m = \frac{60}{T} \cdot 9 \cdot B_t(5) \tag{18}$$

It is important to note that none of the 27 subjects during all spelling tasks ever reached the letters ',' and '␣' (located at the left and the right boundaries of the speller layout), therefore the minimal N = 3 in this study.

The highest theoretically achievable ITR can be roughly calculated as 132.68 bit/min (assuming the already introduced minimal time between two consequent command classifications of 1050 ms):

$$ITR_{max} = \frac{60}{1.05} \cdot \log_2 5 \tag{19}$$

Since this is the ideal situation, the real ITR values will be always below this theoretical value.

III. Results

Fig. 3 shows an example of the copy spelling with the Bremen-BCI system, subject 2 spelled the word "BCI". Fig. 3a presents the EEG signal acquired from site O_z, Fig. 3b - the EEG data used for classification, EEG data replacement with zeros is visible, for simplicity the EEG data prior to each command classification is still shown, Fig. 3c - the changes in the normalized signal powers, and Fig. 3d - the corresponding classifier output. This example was chosen because subject 2 achieved for spelling of this word the outstanding ITR value of 117.39 bit/min. It is important to note, that applying of simple methods like the FFT will not allow the reliable command classification on the basis of such short segments of EEG data, but thanks to the minimum energy combination [11], the signal processing shown in the Fig. 3 is possible.

Three conventional BCI performances, spelling time, accuracy and ITR achieved, for five spelling words over all 27 subjects are summarized in Table I. The format of this Table allows a direct comparison to our previously published results [19]. For every word the minimum (**Min**), maximum (**Max**), mean (**Mean**) and standard deviation (**S.D.**) for each measured variable are obtained by averaging over all subjects who completed the particular spelling task. Accuracies achieved in individual spelling tasks vary considerably, in the range from 54.55% to 100.00%. However, the majority of achieved accuracies are over 92%. The spelling time also varies between subjects, leading to ITRs in the range from 4.61 to 117.39 bit/min with the mean of about 50 bit/min.

Fig. 4a shows individual accuracies and information transfer rates achieved by 27 subjects. Mean information transfer rates across all tasks and subjects was 49.93 ± 26.44. Fig. 4b presents the normalized (by the total number of correct classifications) distribution of the time segment length for all correct classifications. This distribution is independent of the stimulus frequency.

Information transfer rates and accuracies were analyzed to find significant differences within spelling tasks using repeated measures analysis of variance (ANOVA). Data from the 25 subjects who successfully completed all tasks were used for this analysis. The results show that performance measured as ITR across five different spelling tasks differed significantly, $F(4, 96 = 20.373, p < 0.001)$. Post hoc tests revealed that the word "BCI" was significantly

TABLE I

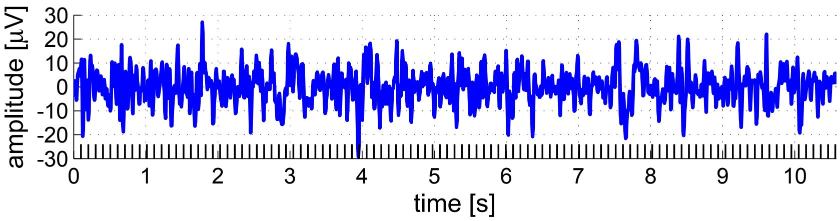
DETAILED RESULTS FOR EACH WORD AND EACH SUBJECT (* REPRESENTS "BCI-ILLITERATES" DURING PREVIOUS EXPERIMENTS).

	Word "BCI"			Word "GEHIRN"			Word "INTERFACE"			Free spelling 1 (5 letters)			Free spelling 2		
Subject	Time	Acc.	ITR	Time	Acc.	ITR	Time	Acc.	ITR	Time	Acc.	ITR	Time	Acc.	ITR
	[s]	[%]	[bpm]	[s]	[%]	[bpm]	[s]	[%]	[bpm]	[s]	[%]	[bpm]	[s]	[%]	[bpm]
1	12.00	100.00	104.47	27.85	100.00	72.06	49.58	93.33	63.93	20.43	100.00	94.52	22.57	100.00	79.39
2	10.68	100.00	117.39	22.87	100.00	87.75	31.91	100.00	95.45	20.74	100.00	100.77	21.24	100.00	97.45
3	12.92	100.00	97.06	28.26	100.00	71.03	38.21	100.00	79.70	21.15	100.00	91.33	20.23	100.00	82.64
4	29.38	100.00	42.68	99.96	95.24	23.21	259.84	84.00	14.90	101.69	86.96	19.04	76.92	100.00	34.16
5*	69.51	100.00	18.04	154.94	88.89	10.43	228.36	87.88	12.96	55.78	100.00	22.48	74.79	90.00	13.26
6	65.12	83.33	21.21	134.60	89.47	13.77	169.86	86.11	17.81	47.98	100.00	23.23	71.52	91.67	17.53
7	26.22	100.00	47.81	49.58	100.00	40.86	71.12	100.00	42.82	41.66	94.12	44.23	9.36	100.00	74.39
8	15.46	100.00	81.12	43.18	100.00	46.47	56.08	100.00	54.31	41.05	100.00	55.21	44.10	100.00	46.95
9	19.72	90.91	56.91	74.57	100.00	27.17	100.17	91.43	34.56	51.92	100.00	37.20	52.43	100.00	50.12
10	13.63	100.00	92.01	32.62	100.00	62.12	48.37	93.55	68.66	32.32	100.00	71.48	18.93	100.00	73.61
11	85.14	78.57	10.96	155.44	89.47	11.11	264.62	90.91	12.24	128.01	94.12	14.40	202.86	84.00	9.47
12	21.85	100.00	57.37	40.14	100.00	50.48	48.16	100.00	63.24	34.46	100.00	58.24	26.94	93.33	59.81
13	11.90	100.00	105.35	27.44	100.00	73.13	38.51	100.00	79.08	31.50	100.00	65.72	16.68	100.00	83.54
14	46.96	100.00	26.70	114.70	87.10	23.42	116.52	86.11	26.27	60.66	100.00	29.54	75.71	94.44	26.11
15*	282.38	54.55	4.61	-	-	-	-	-	-	-	-	-	-	-	-
16	21.96	100.00	57.09	41.05	100.00	48.89	169.94	90.20	28.55	43.90	100.00	31.74	58.83	89.47	38.13
17	12.31	100.00	101.89	25.72	100.00	78.04	37.50	100.00	81.22	24.09	100.00	95.91	21.24	100.00	78.69
18	13.22	100.00	94.83	31.81	94.12	58.50	39.43	100.00	77.25	20.23	100.00	88.58	29.58	100.00	73.87
19	12.92	100.00	97.07	28.56	100.00	70.95	45.52	100.00	66.91	40.74	95.45	61.27	19.22	100.00	94.25
20	46.64	90.48	42.95	-	-	-	-	-	-	-	-	-	-	-	-
21	19.52	90.91	55.51	40.35	100.00	49.74	179.39	79.03	21.70	45.83	88.24	34.81	53.66	77.78	19.85
22	12.31	100.00	101.88	40.24	94.12	46.24	64.82	92.31	61.86	26.63	93.75	65.71	23.17	100.00	71.31
23	34.25	100.00	36.61	146.68	88.57	20.68	441.44	76.62	10.48	64.00	100.00	30.17	165.37	89.19	20.35
24	19.83	100.00	63.24	99.05	83.33	31.41	152.79	77.78	21.97	32.72	94.12	55.77	128.60	65.38	15.48
25	26.95	100.00	46.52	107.89	88.89	15.13	183.16	87.04	25.21	58.22	95.24	39.86	34.47	100.00	51.42
26	25.71	100.00	48.76	67.78	100.00	29.61	150.14	90.38	32.92	66.44	94.44	29.82	24.90	100.00	44.76
27	19.22	100.00	65.25	40.64	94.12	45.34	56.29	100.00	54.11	35.06	100.00	39.74	94.07	92.11	42.25
Min	10.68	54.55	4.61	22.87	83.33	10.43	31.91	76.62	10.48	20.23	86.96	14.40	9.36	65.38	9.74
Max	282.38	100.00	117.39	155.44	100.00	87.75	441.44	100.00	95.45	128.01	100.00	100.77	202.86	100.00	97.45
Mean	36.58	95.88	62.79	67.04	95.73	44.29	121.67	92.27	45.92	45.89	97.46	52.03	55.50	94.69	51.95
S.D.	52.82	10.00	32.57	45.24	5.46	23.14	99.82	7.72	26.39	25.33	3.89	26.50	48.80	8.61	27.65

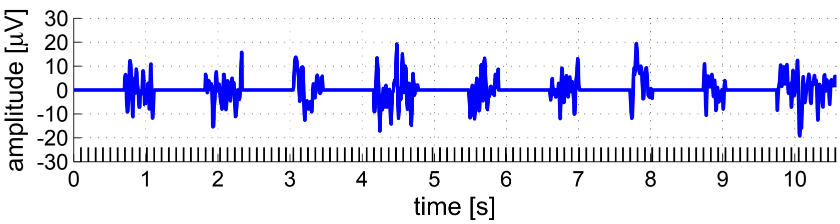
different from all other spelling tasks ($p < 0.002$). In terms of accuracy, also spelling tasks were performed differently $F(4, 96 = 4.70, p < 0.002)$. Post hoc tests show that the word "INTERFACE" differed from others two words: "BCI" $(p < 0.05)$ and free spelling with five letters $(p < 0.05)$.

IV. DISCUSSION

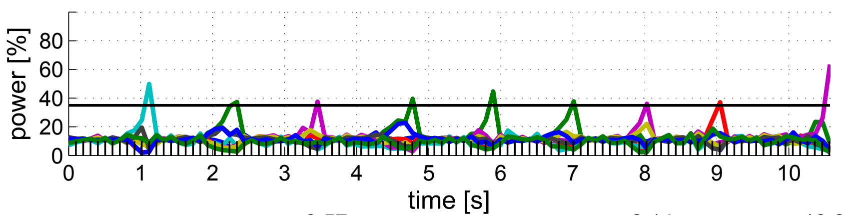
The highest mean ITR for healthy subjects using an SSVEP-based BCI with frequency modulation reported so far is 58 $\pm$ 9.6 bit/min, while the highest ITR reported for one subject was 67 bit/min [8]. The study was conducted with 9 experienced and 3 naive subjects. In other paper from the same research group, *Bin et al.*, reported ITR values of 92.8 $\pm$ 14.1 bit/min for code based modulation [20]. These ITR values are not directly comparable to the study presented here due to different types of modulation. By introducing a higher number of naive subjects, the chance



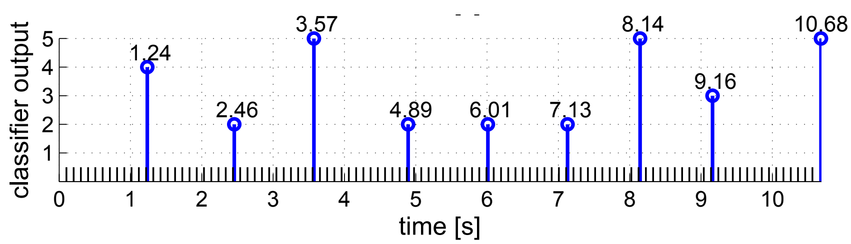
(a) Example of the raw EEG signal acquired from site O_z.



(b) The EEG signal acquired from site O_z used for classification.



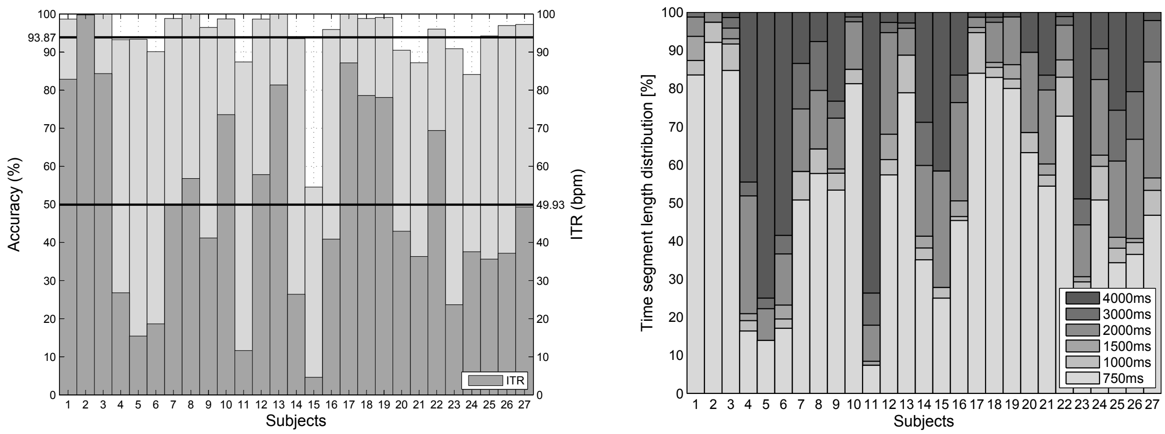
(c) Signal classification on the basis of the fixed threshold $\beta = 0.35$.



(d) Classifier output.

Fig. 3. Spelling with Bremen-BCI on example of the subject 2 and the copy spelling word "BCI". The classifications were performed every 13 samples, these moments are marked with stroke lines on the x axes in the diagrams.

of getting BCI illiterates [21] increases. In previous studies with many naive subjects, we obtained 5 out of 37 [19] and 26 out of 106 (we performed an additional interpretation of the data presented in [22] to determine the number of subjects who spelled at least one of the five words correctly in order to get the percentage of illiterates) subjects as BCI illiterates. In this study, 21 naive subjects participated and we noticed that improving the algorithms could cure the BCI illiteracy for SSVEP based BCIs. Two subjects, who were defined as BCI illiterates in our previous experiments were invited to take part in this study (subject 5 and subject 15). Both of them were able to spell at least one word correctly. Subject 5 even completed all tasks. However, subject 15 aborted the experiment after finishing the first task ("BCI"). Although the task was difficult to complete (with spelling time of 282.38 s and accuracy of



(a) Mean ITRs and accuracies over all subjects

(b) Distribution of the time segment length for all correct classifications

Fig. 4. Results over all 27 subjects. Fig. 4a shows mean individual accuracies and information transfer rates. Fig. 4b presents the distribution of the time segment length for all correct classifications. The time segment of 750ms was most often selected by the system for classification (mean value $53 \pm 25.48\%$). A correlation analysis was performed to determine if a relationship between the information transfer rate and the percentage of time segment length of 750ms exists. The correlation test indicated a strong relationship between the two variables ($r = 0.96, p < 0.001$) by using a Pearson bivariate correlation.

54.55 %), subject 15 was able to write the word. Subject 20 experienced visual fatigue and discomfort, and aborted the experiments, but the ITR achieved by this subject (42.95 bit/min) leads to the conclusion that this subject was able to use the system. For all subjects who wrote at least one word, the mean ITR was 49.93 ± 26.44 bit/min with a mean accuracy of 93.87 ± 9.06 %. Considering only the subjects who finished the complete experiment, these values increased (ITR: 52.02 ± 25.79 bit/min, accuracy: 95.58 ± 4.58 %). In comparison, the mean ITR for naive subjects was 48.44 ± 23.03 bit/min and the mean accuracy was 95.01 ± 4.81 %. Although these values are lower than the reported ITR in [8], [20], these results are promising, because of the high ITR and accuracy for naive subjects. In addition, the highest mean ITR for a single subject is 99.76 bit/min in this experiment (subject 2). If we calculate the mean ITR and accuracy for all subjects, who already participated in previous studies (subject 1-3 and subject 27), it seems, that a much higher ITR is possible when a subject got used to the application (mean ITR: 79.08 ± 21.25 bit/min, mean accuracy: 98.98 ± 1.31 %). The fact that former BCI illiterates are able to complete a spelling task partially validates the successful improvement of all parts of the BCI.

One of the main factors that affects ITR and therefore the communication itself is the speed of the system. An optimal balance of speed and accuracy has to be found in a BCI system for daily use. The impact of the segment length of the data to be analyzed plays an important role in this balancing act. The larger the segment length the better the SSVEP response. *Wang et al.* adapted the segment length to each user to optimize the speed of the system [23]. Based on the suggestion to keep the segment length as low as possible (0.5 s in [24]) and a previous investigation of the accuracy over different time segment lengths [18], the time segment length was continuously adapted to the subject's SSVEP response in this study. The distribution of the time length depicted in Fig. 4b shows that even for the subject with the worst ITR (subject 15) a segment of 750 ms can be used sometimes for a correct classification of the intended frequency. Also, it can be seen that the percentage of the shortest time segment length correlates with the mean ITR of each subject. The larger the percentage of the 750 ms-segment the higher the mean ITR of the system. 92.11 % of true positive classifications were made using a time window length of 750 ms for the best subject (subject 2). In addition, the longest window length was chosen to 2000 ms in 2.63 % of true positive classifications for that subject. The influence of the time segment length is also visible in the results of the ANOVA analyses of the ITR. The word "BCI" was significantly different from the other tasks. The signal processing used in the experiments always starts with a time segment length of 750 ms. If no frequency could be classified, the window was enlarged. A task that uses a higher percentage of a window length of 750 ms results in a significantly higher ITR than a similar task with a different segment length distribution.

Another factor that probably contributes to the high ITR could be the feedback type presented to the user. Although, only subjective reports of the subjects could be obtained, we believe that the continuous feedback (size of the stimuli changed in relation to the SNR signals) had a positive effect on the real-time performance.

V. Conclusion and Future Work

Presented results demonstrated that improvements in the signal processing and new feedback modules of the BCI system constituted the starting basis for achieving ITRs in the order of 100 bit/min. The novel idea of adaptive time segment length and the new type of visual feedback were the keys to achieve high online performance. Further research should identify other factors that can relate with performance, such as, human factors, training procedures, time behavior of the complete system (the user can learn the time latencies and responses of the system), and error recognition and error correction at early stages of the BCI signal processing. In this work, the classification threshold was set to a constant value for all frequencies ($\beta = 0.35$). Further research might also consider a frequency dependent β to take into account the variation of the power of SSVEPs at different frequencies.

Further, water-based or dry EEG-electrodes should be considered to make BCI systems less annoying. Our recent study presents the first results of the evaluation of water-based electrodes in the online BCI experiments with ten healthy subjects [25].

References

[1] G. Dornhege, J. del R. Millan, T. Hinterberger, D. J. McFarland, and K.-R. Müller, *Toward Brain-Computer Interfacing*. MIT Press, 2007.

[2] G. Schalk, "Sensor modalities for brain-computer interfacing," in *Human-Computer Interaction, Part II, HCII 2009, LNCS 5611*, 2009, pp. 616–622.

[3] J. R. Wolpaw, H. Ramoser, D. J. 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] E. C. Lalor, S. P. Kelly, C. Finucane, R. Burke, R. Smith, R. B. Reilly, and G. McDarby, "Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment," *EURASIP J. Appl. Signal Process.*, vol. 19, pp. 3156–3164, 2005.

[5] X. Gao, X. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 137–140, Jun. 2003.

[6] 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, 2008.

[7] F.-B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives," *Prog. Neurobiol.*, vol. 90, pp. 418–438, Feb. 2010.

[8] 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," *J. Neural Eng.*, vol. 6, no. 4, p. 046002, Jun. 2009.

[9] L. J. Trejo, R. Rosipal, and B. Matthews, "Brain-Computer Interfaces for 1-D and 2-D cursor control: Designs using volitional control of the EEG spectrum or steady-state visual evoked potentials," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 14, no. 2, pp. 225–229, June 2006.

[10] K. D. Nielsen, A. F. Cabrera, and O. F. do Nascimento, "EEG based BCI-towards a better control. Brain-computer interface research at aalborg university," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 202–204, June 2006.

[11] O. Friman, I. Volosyak, and A. Gräser, "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.

[12] Z. Wu, Y. Lai, Y. Xia, D. Wu, and D. Yao, "Stimulator selection in SSVEP-based BCI," *Med. Eng. Phys.*, vol. 30, no. 8, pp. 1079–1088, Oct. 2008.

[13] C. S. Herrmann, "Human EEG responses to 1-100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Exp. Brain Res.*, vol. 137, no. 3–4, pp. 346–353, Apr 2001.

[14] M. A. Pastor, J. Artieda, J. Arbizu, M. Valencia, and J. C. Masdeu, "Human cerebral activation during steady-state visual-evoked responses," *J. Neurosci.*, vol. 23, no. 37, pp. 11 621–11 627, Dec. 2003.

[15] I. Volosyak, H. Cecotti, and A. Gräser, "Optimal visual stimuli on LCD screens for SSVEP based Brain-Computer Interfaces," in *Proc. 4th Int. IEEE/EMBS Conference on Neural Engineering NER 09*, May 2009, pp. 447–450.

[16] ——, "Impact of Frequency Selection on LCD Screens for SSVEP Based Brain-Computer Interfaces," in *IWANN 2009, Part I, LNCS 5517*. Springer, 2009, pp. 706–713.

[17] D. Valbuena, I. Sugiarto, and A. Gräser, "Spelling with the Bremen Brain-computer Interface and the Integrated SSVEP Stimulator," in *Proc. 4th Int. Brain-Computer Interface Workshop and Training Course*, Graz, Austria, Sep. 18–21 2008, pp. 291–296.

[18] I. Volosyak, H. Cecotti, and A. Gräser, "Steady-state visual evoked potential response - impact of the time segment length," in *Proc. on the 7th international Conference on Biomedical Engineering BioMed2010, Innsbruck, Austria, February 17 –19*, 2010, pp. 288–292.

[19] I. Volosyak, H. Cecotti, D. Valbuena, and A. Gräser, "Evaluation of the Bremen SSVEP based BCI in real world conditions," in *Proc. IEEE ICORR'09*, Jun. 2009, pp. 322–331.

[20] G. Bin, X. Gao, Y. Wang, B. Hong, and S. Gao, "VEP-based brain-computer interfaces: time, frequency, and code modulations [Research Frontier]," *IEEE Comput. Intelli. Mag.*, vol. 4, no. 4, pp. 22–26, November 2009.

[21] A. Nijholt and D. Tan, "Brain-Computer Interfacing for Intelligent Systems," *IEEE Intell. Syst.*, vol. 23, no. 3, pp. 72–79, 2008.

[22] B. Allison, T. Lüth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Gräser, "BCI Demographics: How many (and what kinds of) people can use an SSVEP BCI?" *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 107–116, Apr. 2010.

[23] 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. 234–239, Jun 2006.

[24] T. Lüth, A. Moltsaar, A. Teymourian, and A. Gräser, "Using individual auto-regression models in an SSVEP-based Brain-Computer Interface," in *Proc. 4th Int. Brain-computer Interface Workshop and Training Course, Graz, Austria*, Sep. 2008, pp. 262–267.

[25] I. Volosyak, D. Valbuena, T. Malechka, J. Peuscher, and A. Gräser, "Brain-Computer Interface using Water-based Electrodes," *J. Neural Eng.*, vol. 7, p. 066007, 2010.