

## Research Article

# A Robust and Self-Paced BCI System Based on a Four Class SSVEP Paradigm: Algorithms and Protocols for a High-Transfer-Rate Direct Brain Communication

Sergio Parini,<sup>1</sup> Luca Maggi,<sup>2</sup> Anna C. Turconi,<sup>3</sup> and Giuseppe Andreoni<sup>2</sup>

<sup>1</sup>Bioengineering Department, Politecnico di Milano University, 20133 Milan, Italy

<sup>2</sup>INDACO Department, Politecnico di Milano University, 20133 Milan, Italy

<sup>3</sup>IRCCS Eugenio Medea "La Nostra Famiglia", 23842 Bosisio Parini, Lecco, Italy

Correspondence should be addressed to Sergio Parini, sergio.parini@polimi.it

Received 11 August 2008; Revised 30 December 2008; Accepted 5 February 2009

Recommended by Li Yuanqing

In this paper, we present, with particular focus on the adopted processing and identification chain and protocol-related solutions, a whole self-paced brain-computer interface system based on a 4-class steady-state visual evoked potentials (SSVEPs) paradigm. The proposed system incorporates an automated spatial filtering technique centred on the common spatial patterns (CSPs) method, an autoscaled and effective signal features extraction which is used for providing an unsupervised biofeedback, and a robust self-paced classifier based on the discriminant analysis theory. The adopted operating protocol is structured in a screening, training, and testing phase aimed at collecting user-specific information regarding best stimulation frequencies, optimal sources identification, and overall system processing chain calibration in only a few minutes. The system, validated on 11 healthy/pathologic subjects, has proven to be reliable in terms of achievable communication speed (up to 70 bit/min) and very robust to false positive identifications.

Copyright © 2009 Sergio Parini et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

A brain-computer interface (BCI) is a system dedicated at providing its users with a new and alternative communication channel totally independent from the traditional output pathways of the nervous system such as peripheral nerves and muscles [1]. A BCI system achieves this goal by directly interfacing the cerebral activity, being it evoked or self-induced, with a common personal computer which nowadays represents a powerful and affordable platform for productivity, entertainment, worldwide communication, and remote control. Though strongly promoted and encouraged by military and entertainment research, its primal and main application is undoubtedly assistive technology for people affected by severe motor disabilities. In recent years, the research in the BCI field has grown rapidly, showing renewed interest and demonstrating how this communication system is in principle feasible. Nonetheless, present day solutions still have shortcomings that prevent their widespread application: limited information transfer-

rate, high susceptibility to false positives, intense and wearing calibration sessions, and sensors intrusively are only a few of the major issues of current BCI systems.

In this paper we present, focusing on the adopted algorithmic and protocol-related solutions, a whole brain-computer interface system based on the steady-state visual evoked potentials (SSVEPs) paradigm. Stimulation at a certain frequency leads to oscillations at the same frequency and harmonics/subharmonics of the stimulation frequency [2]. In a BCI system, SSVEPs are used by simultaneously displaying several stimuli flickering at different frequencies. A user is able to select one specific stimulus by focusing on it consequently leading to an increased amplitude localized on those frequency bands related to the flickering frequency of the stimulus itself.

By considering other systems employing this particular BCI paradigm [2–5], the obtained bit-rates can be considered high when compared to other kinds of BCI systems (ranging from 30 to 68 bit/min depending on the protocol, the stimulation device, the number of classes, and the subject).

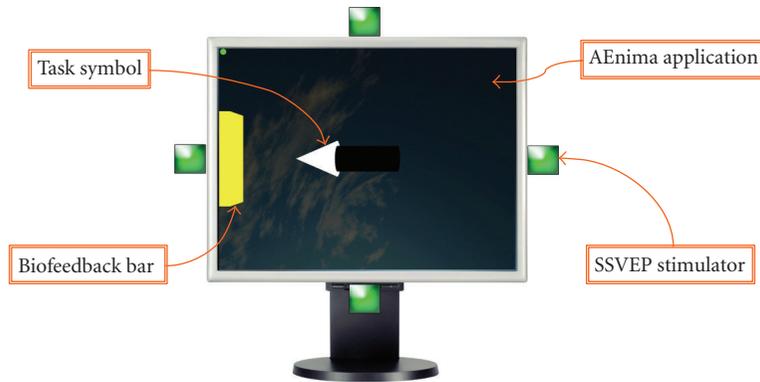


FIGURE 1: The stimulators placement and the graphic user interface layout.

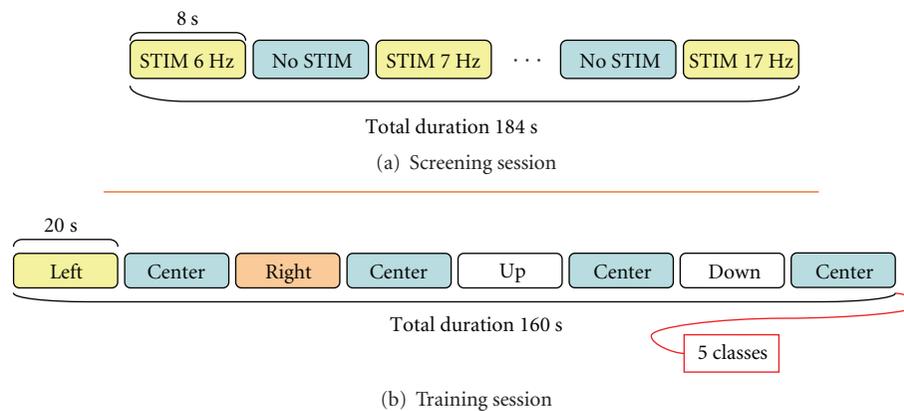


FIGURE 2: Timings and structures of the adopted (a) screening and (b) training protocols.

In spite of the great advantage in terms of speed, SSVEP-based systems have been considered as “second class BCIs” over the years since a residual control of the gaze direction is needed (dependent BCIs). The lack of interest in SSVEP is reflected in the roughness of the current systems which implement basic mechanisms of adaptation to the user specific characteristics. Recent works [6] had renewed the interest in SSVEP-based systems: “*Many other important questions involving practical long-term use of SSVEP BCIs have not been addressed. Training with SSVEP BCIs can improve performance ... the labels “dependent” and “independent” might be best regarded not as absolutes, but endpoints of a continuum.*”

The main objectives of this study are guarantee a self-paced and asynchronous control maximizing communication performances in terms of attainable maximum bit-rate and robustness to false positives, provide a totally unsupervised biofeedback for an optimal mutual man-machine learning, and minimize the whole calibration phase duration.

## 2. Materials & Methods

The proposed SSVEP-based BCI system makes use of the BCI++ Framework [7, 8] as a technological platform for data

acquisition, real-time algorithms management, protocols development, stimulation, and user interfacing. The framework structure can be summed up in two main modules able to communicate remotely by means of a TCP/IP socket based layer.

- (i) *HIM Module*: the hardware interface module dedicated to the acquisition, storage, and visualization of the signal, the communication with the user interface, and the real-time execution of algorithms developed using C/C++ or Matlab environment;
- (ii) *AEnima Module*: the user interface module studied in order to simplify the implementation of new operating protocol for BCI-based user applications. This module was written using an OpenGL/Direct3D-based graphical engine in order to provide a more realistic and challenging experience to the user.

The visual stimulation system consisted of four cubic spotlights with sides of 3 cm; each cube was mounted on a specific side of a standard 19” LCD monitor, thus allowing the user to ideally associate each light source to a 2-dimensional direction: up, down, left, or right (Figure 1). Each cube included a high-efficiency S-Flux green ( $\lambda = 500$  nm) LED. In order to avoid direct exposition to the light and to diffuse

the beam in a more efficient way, a matt film was put on the exposed face of the cube acting as a filter.

*2.1. Operating Protocol.* The adopted operating protocol aimed at gathering and confirming user-specific parameters for an optimal configuration of the processing and identification chain. The aim is that of overcoming potential intersubjects variability issues and guaranteeing the man-machine mutual-adaptation loop to start from the most reliable basis. The operating protocol consisted of the following sessions.

- (i) Screening Session: to identify the most suitable stimulation frequencies for the subject.
- (ii) Training Session: to configure and train the processing and identification chain parameters.
- (iii) Testing Session: to validate and confirm the configuration parameters.

All the described sessions were guided by means of a specific graphic user interface developed using the AEnima module which also managed the stimulation device.

*2.1.1. The Screening Session.* This session was structured in order to gather data useful for identifying the four most effective stimulation frequencies for the specific user: the amplitude of the SSVEP is not the same for different stimulation frequencies or different subjects [9]. The protocol consisted in two different phases:

- (i) a stimulation phase: the user had to focus on a single light source for eight seconds;
- (ii) a resting phase: the stimulation was turned off, and the user had to focus on the center of the screen for eight seconds.

Each resting phase was followed by a stimulation phase at increased frequency resulting in a complete check of a predefined frequency range with a 1 Hz step (Figure 2(a)). The whole screening session lasted about 3 minutes (184 seconds).

According to [5], an optimal signal-to-noise ratio (SNR) of the steady-state visual evoked response can be usually achieved by stimulating in the 5–30 Hz range, with an SNR peak at 15 Hz which decreases at higher stimulation rate [9]. Moreover, according to [10], the late theta, the whole alpha band (6–13 Hz), and the early beta band (13–17 Hz) usually show better stimulus-related responses, while higher stimulation rate has proven to be reliable only when inducing a specific resonance phenomenon indicating a selective frequency preference of the neural oscillators of the subject. In this study, we decided to limit the inspected frequency range in the 6–17 Hz interval taking into account the following purposes:

- (i) minimize the screening session duration;
- (ii) attain to the 5–30 Hz optimal range;

(iii) include the theoretically optimal 15 Hz stimulation rate and its adjacent frequencies;

(iv) minimize the inclusion of the mid-late beta (15–25 Hz) band because of its higher risk of inducing photoepileptic seizures [11];

A similar frequency range has been successfully adopted by other research groups [12]. In addition, the choice of stimulation frequencies in an SSVEP-based BCI application must ensure that the responses are as unique as possible. Thus, the stimulation frequencies were chosen in order to be neither harmonics nor subharmonics of each other [2].

The collected screening dataset was fed into a specific processing tool which calculated, for each common-average referenced electrode (each electrode referenced to the linked mastoids, see Figure 5), the joint time-frequency analysis (JTFA) by means of a short-time Fourier transform with a 4-seconds length (0.25 Hz frequency resolution) Kaiser window ( $\beta = 0.5$ ) and 75% overlap. The JTFA parameters were chosen in order to guarantee a reasonable frequency resolution (0.25 Hz) limiting the time resolution (0.25 seconds) thus reducing computational demand. The choice of a Kaiser window was due to its desirable characteristics of minimizing both main lobe duration and side-lobes area thus reducing spectral leakage phenomena [13]. The JTFA results were plotted using a colormap-based graph and presented to the operator for visual inspection (Figure 3). Using the JTFA and according to the structure of the screening session (Figure 2), the operator could identify the response of each stimulation frequency (fundamental marker) and, consequently, the related response at higher or lower harmonics. For example, in the particular case presented in Figure 3, it is possible to notice a very weak fundamental response of the 6 and 7 Hz stimulations though showing a power increase of the harmonics of the stimulation rate. Higher stimulation frequencies clearly show both fundamental and harmonic response thus suggesting a more reliable visual evoked potential.

*2.1.2. The Training Session.* This session aimed at collecting data useful for defining optimal processing and identification parameters for the specific user. During this session, all the four light sources were activated, and their flashing rates were set to the subject specific relevant stimulation frequencies identified at the end of the screening session. The protocol structure consisted in a guided sequence of stimulus gazing tasks; a symbol displayed on the LCD monitor guided the user to concentrate on a specific light source (left/right/up/down cue-symbol) or on the screen centre (dot-symbol) for 20 seconds following the sequence shown in Figure 2(b).

*2.1.3. The Testing Session.* This session was used to verify and validate the performance of the system configured using the parameters defined in the previous sessions.

The protocol structure consisted in a guided sequence completion task. The user was asked to focus his/her attention on a particular light source, while the signal

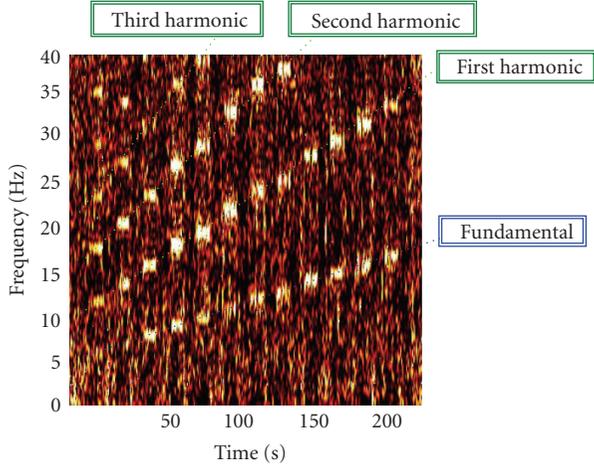


FIGURE 3: An example of a joint time-frequency analysis (JTFA) calculated on a screening dataset varying the stimulation frequency in the 6–17 Hz range.

was processed and identified continuously by the self-paced online translation algorithm; the switching from one stimulus to another occurred only after the system identified a command related to the actual target source.

The sequence consisted of eight stimulus-specific tasks (two tasks per target) in random sequence: the classifier was activated four times per second, and a classification was considered reliable only when five consecutive and coherent classifications were produced. After each task, stimulation was temporarily stopped for three seconds before switching to the next task.

During this session, the unsupervised biofeedback was continuously provided to the user in order to help him/her in finding the best and most effective stimulus fixation strategy.

**2.2. The Processing and Identification Chain.** The processing and identification chain adopted for the real-time, self-paced translation algorithm is summarized in Figure 4. Each single block of the flow diagram is here described in details.

**2.2.1. Pre-processing and Data Windowing.** The 8-channel data acquired by the input device were continuously notch filtered on AC frequency using an elliptic filter with approximately 60 dB stopband attenuation. A further wide passband 5–45 Hz filter (4th-order FIR) was applied in order to reduce EOG/EMG artefacts. Preprocessed data were windowed using a 3-second length analysis window.

**2.2.2. Spatial Filtering and Channel Combining.** Windowed data were spatially filtered using a single linear combiner with channel-specific weights. The aim of this processing block was that of optimizing the information contained in the 8-channel data in order to maximize the stimulus-related response signal-to-noise ratio with respect to basal activity.

Based on the assumption that an optimal spatial filtering was independent from the specific stimulation frequency, only a single linear combiner was adopted thus resulting

in an 8-to-1 channel data compression. The spatial filter estimation was based on the training dataset and carried out using both a manual and an automatic approach.

The automatic approach was based on the common spatial patterns (CSPs) [14] method which has already proven to give reliable and effective results when used with other motor-imagery-based BCI applications [15–17]. Starting from two distributions in a  $C$ -dimensional space, being  $C$  the number of acquired channels, the CSP algorithm finds projections that maximize variance for one class while minimizing the variance of the other class. In the specific case of the SSVEP, the fixation of a light source flashing at a predefined frequency causes the brain to start “tuning” on that firing frequency thus producing a quasiperiodic response: by filtering the signal with a narrow bandpass filter centred on the stimulation frequency, this stimulus-related variation results in a higher variance signal with respect to the nonstimulus condition. From this point of view, the SSVEP signal seems well suited for the application of the CSP method intended as a technique for maximizing the SNR of the visual-evoked response against the nonstimulus condition.

According to the previously stated assumption, in order to exploit the information contained in the whole training dataset, the CSP estimation was applied on a repacked version of the training dataset consisting of 2 classes *stimulus* and *nonstimulus*. Each stimulus related portion of the dataset was narrow bandpass filtered ( $\pm 0.2$  Hz, stopband attenuation 30 dB) on the stimulation frequency and on its first and second harmonics. Similarly, each nonstimulus-related (NULL) portion was narrow bandpass filtered on the same frequency of the previous stimulation phase and on its harmonics. The contribution of each harmonics was considered additive. In this way, the 5 classes dataset (left, right, up, down, null) was repacked and consequently relabelled in a 2 classes dataset containing stimulus related and nonstimulus related trials. Calling  $X$  the repacked dataset, let  $\Omega_S$  and  $\Omega_{NS}$  be, respectively, the pooled estimates of the covariance matrices for the *stimulus* and *nonstimulus* related classes, as follows:

$$\Omega_{\text{class}} = \frac{1}{N_{\text{class}}} \sum_{i=1}^{N_{\text{class}}} X_i X_i^T \quad \text{with } (\text{class} \in \{S, NS\}), \quad (1)$$

where  $N_{\text{class}}$  is the number of trials for class, and  $X_i$  is the matrix (in the form *(channels, sample)*) containing the portion of data corresponding to the  $i$ th trial. The two covariance matrices,  $\Omega_S$  and  $\Omega_{NS}$ , are simultaneously diagonalized such that the eigenvalues sum to one. This is achieved by calculating the generalized eigenvectors  $V$  as follows:

$$\Omega_S V = (\Omega_S + \Omega_{NS}) V L, \quad (2)$$

where  $L$  is the diagonal matrix containing the generalized eigenvalues, and the column vectors of  $V$  are the filters for the CSP projections. Theoretically, in the specific case of SSVEP, the best contrast is provided only by those eigenvectors corresponding to eigenvalues close to 1 (large variance for

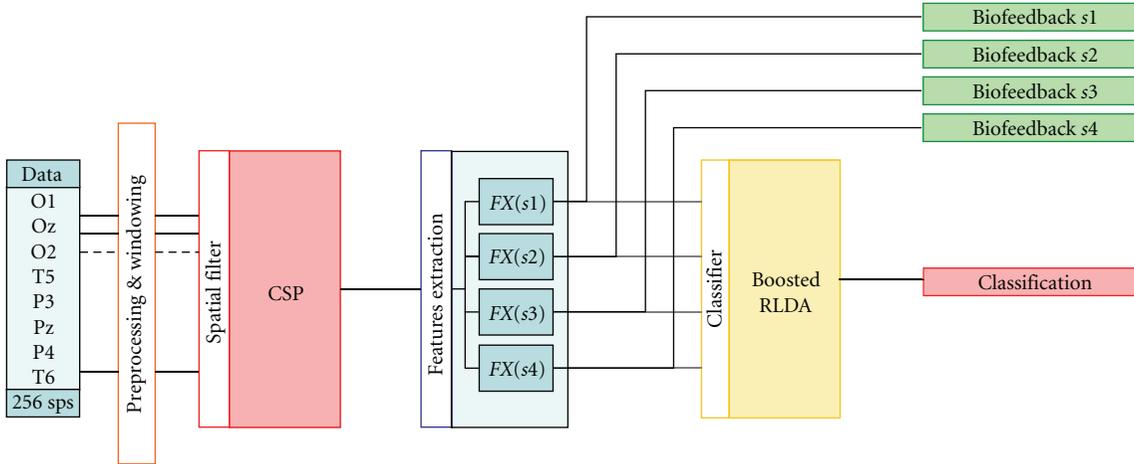


FIGURE 4: Conceptual scheme of the processing and identification chain for each of the four stimulation frequencies ( $s_1, s_2, s_3, s_4$ ).

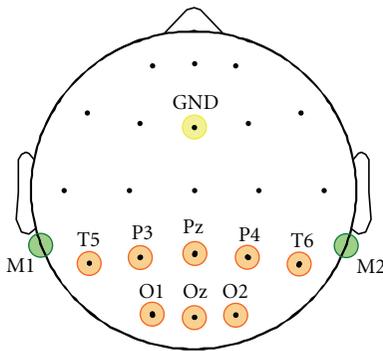


FIGURE 5: The adopted electrodes montage includes orange highlighted channels  $\{T_5, P_3, P_z, P_4, T_6, O_1, O_z, O_2\}$  using a linked mastoids reference and a ground channel placed in GND position.

class S and small variance for class NS). Thus, in this study, we chose to consider only the filter corresponding to the highest eigenvalue. According to [15], useful information is also contained by those projections corresponding to eigenvalues close to the highest; in this specific work, we decided not to consider this information in order to minimize the feature vector dimensionality and maximize classifier robustness with small training dataset.

The manual approach relied on a software tool developed for the selection of channels combinations. The operator was asked to specify the type of contribution related to each electrode: *positive-contribution*, *negative-contribution*, and *null-contribution*. The resulting data channel was calculated by subtracting the mean of the *negative-contribution* electrodes from the mean of the *positive-contribution* electrodes, ignoring the *null-contribution* channels. The manual selection of the channels combination was aided by providing the operator with a binary class separability index (Section 2.2.3) calculated, for each stimulation frequency, using the features extracted from the idle portion and the frequency-specific stimulation portion of the training dataset. This specific approach was considered in order to give the operator

the possibility to evaluate, aided by the CSP mapping, the exclusion of channels showing low discriminative power (filter weight close to 0), thus optimizing the subject specific electrodes montage.

**2.2.3. Features Extraction and Evaluation.** The signal feature adopted in this study is an adaptation of the stimulus-locked average frequently used for the extraction of evoked potential. In this case, no flash-locked triggering was used but instead we assumed that each event-related response has duration equal to the specific sth stimulation period  $p_s$ . According to this, the windowed and spatially combined signal  $X_{win} = [x_1, x_2, \dots, x_N]$  was divided into a number  $N_{SW}$  of subwindows of length  $p_s$  and averaged synchronously, thus

$$B_s = [\beta_s(1), \beta_s(2), \dots, \beta_s(p_s)], \quad (3)$$

where

$$\beta_s(i) = \frac{1}{N_{SW}} \cdot \sum_{j=1}^{j=N_{SW}} X_{win}(i + (j-1) \cdot p_s), \quad (4)$$

being  $N_{SW} = \frac{N}{p_s}$ ,

where the signal components, synchronous with the stimulation frequency, were consequently enhanced while the other components were reduced.

An amplitude estimation of the SSVEP was obtained by means of the following equation:

$$FX(s) = \frac{\sigma(B_s)}{\sigma(X_{win})}, \quad (5)$$

where  $\sigma(B_s)$  is the standard deviation of the average (time-locked to the sth stimulus),  $\sigma(X_{win})$  is the standard deviation calculated using the whole data window, and  $FX(s)$  is the resulting normalized feature, estimation of the response to the sth stimulus. The resulting feature is a numerical value which theoretically ranges from zero, in the case of total disruptive interference, to one, when the signal has

only one periodic component with the same period of the stimulation (constructive interference). The main advantage of this property is that the resulting values do not need to be scaled on the basis of a previous training session thus allowing the operator to provide the user with a consistent biofeedback right after the initial screening session. Only one feature per stimulus was extracted thus resulting in a final features vector of 4 elements.

Once the features were extracted, an estimation of the discriminative power of the system was provided in the form of a Bhattacharyya distance calculated between each stimulation-related portion of the features-set and its consequent nonstimulation portion according to the training dataset structure (Figure 2(b)).

The Bhattacharyya bound is a special case of the Chernoff bound [16, 17]. Often, the Bhattacharyya bound is used instead of the Chernoff bound since empirical evidence indicates it to be optimal when the majority of the separation comes from a difference in classes means [18]. Recently, it has also been shown that it is feasible to predict the classification error accurately using the Bhattacharyya distance [19, 20]. The expression for the Bhattacharyya distance is

$$\begin{aligned} \bar{\mu}(c) = & \frac{1}{8} (\bar{M}_{NS} - \bar{M}_c)^T \left[ \frac{\bar{S}_c + \bar{S}_{NS}}{2} \right]^{-1} (\bar{M}_{NS} - \bar{M}_c) \\ & + \frac{1}{2} \ln \frac{[(\bar{S}_c + \bar{S}_{NS})/2]}{\sqrt{|\bar{S}_c| |\bar{S}_{NS}|}} \end{aligned} \quad (6)$$

with  $(c \in \{\text{left, right, up, down}\})$ ,

where  $\bar{M}_c$  and  $\bar{S}_c$  are, respectively, the vectors of means and variances of the class  $c$  of the feature-set, while  $\bar{M}_{NS}$  and  $\bar{S}_{NS}$  are, respectively, the vectors of means and variances of the nonstimulus class of the feature-set.  $\bar{\mu}(c)$  is a vector of 4 elements containing the discriminative power of each feature with respect to class  $c$ . An effective configuration of the system would present, for each class  $c$ , a high discriminative power of the features related to the class specific stimulus while indices close to zero for the other features.

**2.2.4. The Classifier.** The previously described features set was used to train a 5-class (left, right, up, down, null) supervised classifier based on the Regularized Linear Discriminant Analysis method [18, 21]. Linear Discriminant Analysis (LDA) is a well-known classification method that projects high-dimensional data onto a low-dimensional space where the data are reshaped to maximize class separability. The optimal projection or transformation in classical LDA is obtained by minimizing the within-class distance and maximizing the between-class distance simultaneously, thus achieving maximum discrimination.

In this study, we adopted a regularized version of the discriminant analysis method using a double parameter  $(\lambda, \gamma)$  model selection scheme as described in [21]. The proposed identification algorithm also included a boosting technique based on the Freund and Schapire method as described in [22]; classifiers are constructed on weighted versions of the training set, which are dependent on previous

classification results. Initially, all objects have equal weights, and the first classifier is constructed on this dataset. Then, weights are changed according to the performance of the classifier. Erroneously classified objects get larger weights, and the next classifier is boosted on the reweighted training set. In this way, a sequence of classifiers is obtained, which are then combined by a weighted majority vote in the final decision.

### 3. Results

The system was tested on two different control groups made of volunteers without previous BCI experiences. The first control group (G1) was composed of 7 healthy subjects, aged from 18 to 50 years. The second control group (G2) was composed of 4 pathological subjects, aged from 10 to 30, and affected by Duchenne muscular dystrophy at different stages. Every subject suffering from defects of vision performed the whole experimental session wearing appropriate corrective lens. During acquisitions, the subject sat on a comfortable chair at a distance of about 80 cm from an LCD monitor in a noise-controlled and slightly dim room.

The EEG signal was acquired at 256 samples per second by means of an 8-channel wearable system [23] developed by Sensibilab and WoWS!-Lab (Politecnico di Milano University, Italy): the communication with the HIM Module was achieved through a Bluetooth (BT) wireless connection. In this study, we used a g.EEGcap (g.Tec GmbH, AU, <http://www.gtec.at/>) with a set of eight gold-plated electrodes placed over the occipital area {T5,P3,Pz,P4,T5,O1,Oz,O2} with a linked-mastoids reference as illustrated in Figure 5.

The adopted experimental protocol was the same for both control groups and, according to the previously described protocol structures, it consisted in the following:

- (i) a screening session (duration: 184 seconds);
- (ii) a training session (duration: 160 seconds);
- (iii) a testing session (duration: depending on user's performances).

In Table 1, we report, for each subject of group G1 and G2, the global misclassification rate, the typical delay time, and the resulting bit-rate obtained using the CSP-based spatial filtering approach. Though each subject performed the whole experimental session using a system configured on a 3-second long analysis window, results obtained by simulating the testing session using an analysis window of, respectively, 2 and 4 seconds length are also presented.

The misclassification rate was calculated taking into account the self-paced nature of the identification chain, thus increasing the errors count every time the system identified a control state (left, right, up, down) different from the actual control task. The typical delay time consisted in the mean time interval between the control task trigger and the first correct identification.

The bit-rate was calculated according to the popular definition proposed by Wolpaw and based on Shannon's

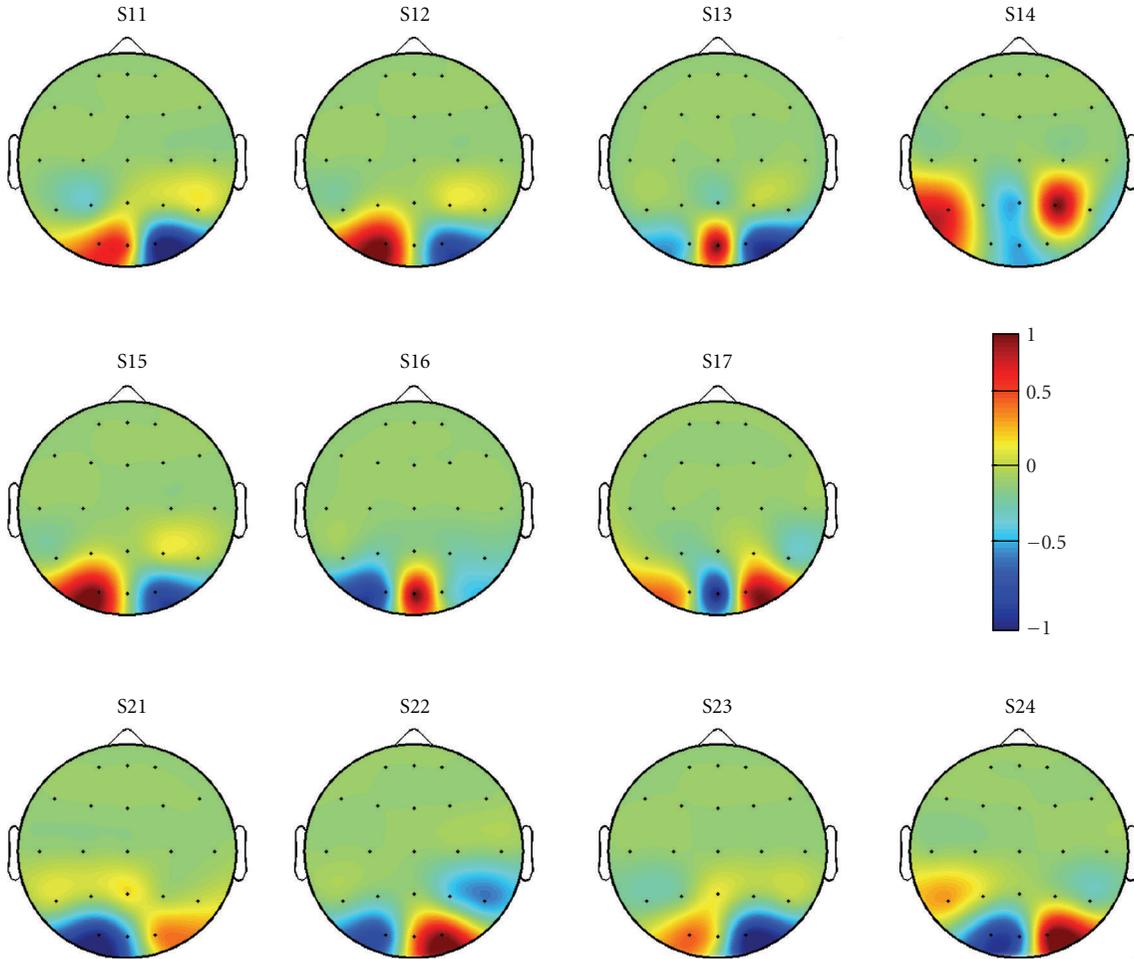


FIGURE 6: Normalized weights associated to each electrode of the considered electrodes montage calculated using the Common Spatial Patterns method. Most of the subjects show a dominant activation localized over the O1-Oz-O2 electrode positions and, according to the weight signs, the optimal filter is usually a bipolar derivation possibly reinforced by parietotemporal activations.

definition of information rate for noisy channels [24] as follows:

$$BR = V \cdot R, \quad (7)$$

where

$$R = \log_2 N + P \cdot \log_2 P + (1 - P) \cdot \log_2 \frac{1 - P}{N - 1}, \quad (8)$$

where BR is the bit-rate in bit/min,  $V$  is the classification speed in symbols/min, and  $R$  is the information carried by one symbol in bit/symbol. The information carrier  $R$  was calculated considering  $N$  as the number of actual control states ( $N = 4$ ) and  $P$  as one minus the global error rate. This definition has widely been used by other research groups (see [24]) in order to evaluate and quantify BCI communication performances, and it is consequently adopted in this study for the sake of comparability.

The error rate ranges from 0 to approximately 7%. Considering a window length of 3 and 4 seconds, the 80% of the analyzed subjects obtained an error rate comprised

from 0 to 1% showing an almost perfect robustness of the system to false positives. The use of a window length of 2 seconds increases the error rate yet showing an average across subjects of about 2.5% which, together with the reduction of the typical delay time, reflects the dramatic gain in terms of bit-rate achieved by the system using small data windows. For every subject, with the only exception of subject S13, the bit-rate ranges from 35 to 70 bit/min with an average across subjects of 51.47 bit/min, a minimum of 17 bit/min, and a maximum of almost 70 bit/min. All the subjects completed the whole testing protocol in less than 80 seconds. The minimum duration of the testing session across all subjects was achieved by subject S12 with an actual session duration (including the pause of three seconds after each task, Section 2.1.3 of approximately 50.85 seconds.

All the results are calculated using the CSP-based spatial filtering approach. In Figure 6, we report, for each subject of group G1 and G2, the normalized weight associated to each electrode of the considered montage. Most of the subjects show a dominant activation localized over the O1-Oz-O2 electrode positions and, according to the weight

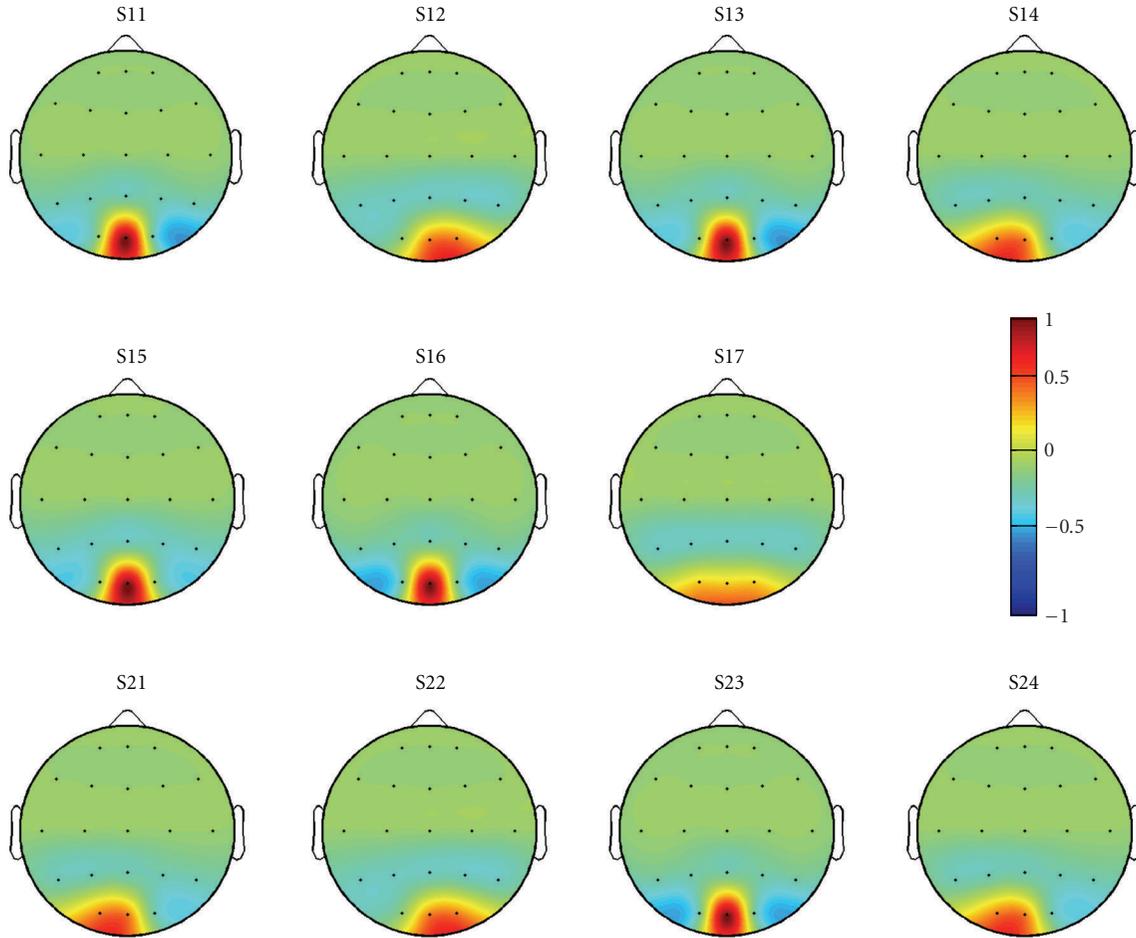


FIGURE 7: Weights associated to each electrode of the considered electrodes montage using the manual channels combining approach described in Section 2.2.2. Positive weights refer to *positive-contribution* electrodes, while negative weights refer to *negative-contribution* electrodes.

TABLE 1: Results using the CSP method for automatic spatial filtering as described in Section 2.2.2. Error rate, typical delay time, and bit-rate calculated on the basis of the testing session performed by each subject of group G1 and G2 using different data analysis window lengths. For each group, the average values across subject are presented (MEAN). A global results calculated on every subject from both control groups is also provided (G1 + G2).

ID	Data window length									
	2 seconds			3 seconds			4 seconds			
	ERR [%]	DLY [sec]	BR [bit/min]	ERR [%]	DLY [sec]	BR [bit/min]	ERR [%]	DLY [sec]	BR [bit/min]	
G1	S11	0.57	2.25	<b>51.83</b>	0.00	2.84	<b>42.26</b>	0.00	3.29	<b>36.53</b>
	S12	0.62	1.68	<b>69.31</b>	0.00	2.30	<b>52.16</b>	0.62	2.58	<b>44.98</b>
	S13	7.03	5.39	<b>16.93</b>	6.83	11.33	<b>8.11</b>	6.21	11.93	<b>7.88</b>
	S14	0.00	3.25	<b>36.88</b>	0.00	3.63	<b>33.07</b>	0.00	3.97	<b>30.21</b>
	S15	3.01	2.49	<b>42.37</b>	0.00	3.18	<b>37.79</b>	0.00	3.66	<b>32.79</b>
	S16	2.26	1.77	<b>61.32</b>	0.00	2.33	<b>51.46</b>	0.00	2.85	<b>42.14</b>
	S17	4.97	1.71	<b>57.24</b>	4.42	2.46	<b>40.62</b>	2.76	3.02	<b>35.25</b>
<b>MEAN</b>	2.64	2.65	<b>47.98</b>	1.61	4.01	<b>37.92</b>	1.37	4.47	<b>32.82</b>	
G2	S21	0.00	1.77	<b>67.81</b>	0.00	2.18	<b>55.15</b>	0.00	2.55	<b>47.04</b>
	S22	5.42	2.05	<b>47.10</b>	3.01	2.71	<b>38.95</b>	4.82	3.21	<b>30.78</b>
	S23	1.18	1.89	<b>59.81</b>	0.59	2.49	<b>46.74</b>	0.00	2.94	<b>40.80</b>
	S24	2.40	1.94	<b>55.60</b>	0.00	2.38	<b>50.44</b>	0.00	2.93	<b>41.01</b>
<b>MEAN</b>	2.25	1.91	<b>57.58</b>	0.90	2.44	<b>47.82</b>	1.20	2.91	<b>39.91</b>	
<b>G1 + G2</b>	2.50	2.38	<b>51.47</b>	1.35	3.44	<b>41.52</b>	1.31	3.90	<b>35.40</b>	

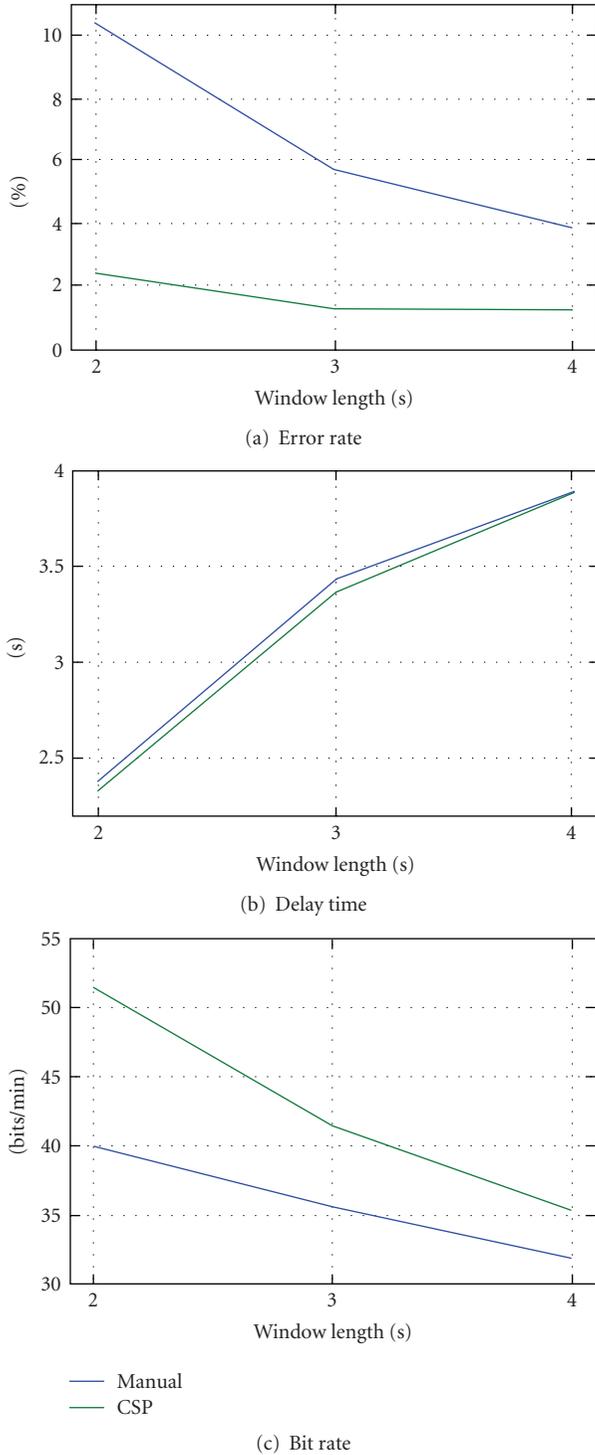


FIGURE 8: Average results across all subjects (G1 + G2) obtained using the manual (blue line) and CSP-based (green line) methods with different lengths of the analysis window; (a) error rate, (b) delay time, and (c) bit-rate calculated using Wolpaw’s definition.

signs, the optimal filter is usually a bipolar derivation (e.g., O1-O2 for S15 and O1-Oz for S16) possibly reinforced by parietotemporal activations (e.g., subject S24). Such distributions suggest that a subject specific optimization of the electrodes montage could provide interesting results

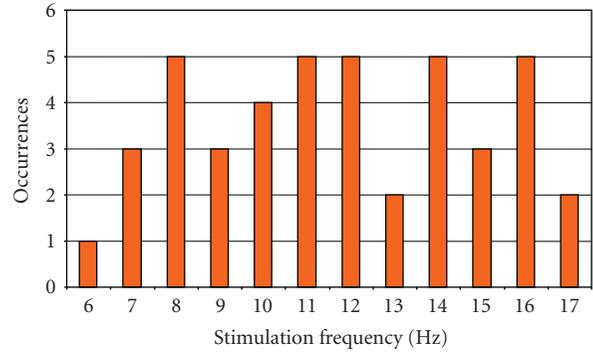


FIGURE 9: Occurrences of each stimulation frequency of the screened 6–17 Hz range calculated on both G1 and G2 control groups.

using only few electrodes, thus a future investigation in this terms is mandatory.

For sake of completeness, in Table 2 we also provide results obtained by means of the standard manual channel combining approach. By comparing the bit-rates obtained using the manual and the CSP method for each subject and for each length of the analysis window, the advantages of the proposed spatial filtering method are clear: the CSP method allows for a considerably lower error rate without worsening the delay time of the SSVEP response identification, thus guaranteeing a huge increase in the achievable information transfer rate. In Figure 7, we also report, for each subject, the normalized weight associated with each electrode of the considered montage using the manual combining approach described in Section 2.2.2.

In Figure 8, for each spatial combining method, we provide (a) the average error rate, (b) delay time, and (c) bit-rate calculated across all the subjects from both control groups (G1 and G2) using different lengths of the data analysis window (2, 3, and 4 seconds). It is possible to notice how the adoption of the CSP method dramatically reduces the error rate yet retaining a comparable delay time of the SSVEP response when compared to the manual channel selection approach. The gap in the resulting error rate is thus reflected in the final bit-rate estimation. In Figure 8(a), the error rate obtained using the CSP method is almost constant between each considered window length. Consequently, the bit-rate gap progressively reduces by increasing the data analysis window length; the proposed CSP-based method is a valuable solution for increasing the identification robustness using small data portion thus offering faster response time and higher information transfer-rates.

In Figure 9, a histogram calculated on the basis of the occurrences of each stimulation frequency in the predefined range (6–17 Hz) is presented. Though beyond the scope of our work, we could notice that, in spite of a low response to the 6 Hz stimulation, no clearly predominant stimulation pattern can be inferred from the results obtained on the considered control groups. This could be seen as a further confirmation of the importance of the identification of subject-specific relevant stimulation frequencies.

TABLE 2: Results using the manual channels combining approach described in Section 2.2.2. Error rate, typical delay time, and bit-rate calculated on the basis of the testing session performed by each subject of group G1 and G2 using different data analysis window lengths. For each group, the average values across subject are presented (MEAN). A global results calculated on every subject from both control groups is also provided (G1 + G2).

ID	Data window length									
	2 seconds			3 seconds			4 seconds			
	ERR [%]	DLY [sec]	BR [bit/min]	ERR [%]	DLY [sec]	BR [bit/min]	ERR [%]	DLY [sec]	BR [bit/min]	
G1	S11	17.05	1.62	<b>39.64</b>	9.66	2.81	<b>29.67</b>	5.11	3.28	<b>29.80</b>
	S12	0.00	1.99	<b>60.35</b>	0.00	2.27	<b>52.87</b>	0.62	2.71	<b>42.90</b>
	S13	14.34	5.83	<b>12.14</b>	10.83	11.21	<b>7.14</b>	10.39	11.77	<b>6.90</b>
	S14	0.00	2.75	<b>43.57</b>	0.00	3.44	<b>34.87</b>	0.00	3.96	<b>30.33</b>
	S15	3.61	2.04	<b>50.66</b>	1.20	2.57	<b>44.11</b>	0.00	3.14	<b>38.16</b>
	S16	10.73	1.74	<b>46.18</b>	2.26	2.55	<b>42.54</b>	0.56	3.24	<b>35.96</b>
	S17	25.97	1.62	<b>28.21</b>	13.81	2.25	<b>32.10</b>	8.84	2.68	<b>31.95</b>
<b>MEAN</b>	10.24	2.51	<b>40.11</b>	5.39	3.87	<b>34.76</b>	3.65	4.40	<b>30.86</b>	
G2	S21	4,80	1,90	<b>51,98</b>	4,10	2,22	<b>45,62</b>	3,12	2,80	<b>37,50</b>
	S22	12,30	2,20	<b>34,55</b>	7,12	2,81	<b>32,38</b>	5,50	3,25	<b>29,64</b>
	S23	10,06	2,24	<b>36,72</b>	6,51	2,58	<b>36,01</b>	7,10	3,02	<b>30,16</b>
	S24	16,17	1,82	<b>36,52</b>	8,38	2,47	<b>35,22</b>	1,80	2,94	<b>37,57</b>
	<b>MEAN</b>	10,83	2,04	<b>39,94</b>	6,53	2,52	<b>37,31</b>	4,38	3,00	<b>33,72</b>
<b>G1 + G2</b>	10.45	2.33	<b>40.04</b>	5.80	3.37	<b>35.68</b>	3.91	3.89	<b>31.90</b>	

#### 4. Discussion and Conclusion

In this work, novel methods dedicated to the SSVEP paradigm are proposed in order to guarantee user specific customization and increase man-machine interaction. Figure 9 supports the introduction of a method for the identification of the subject specific frequencies. The proposed method requires a session of only 184 seconds. The analysis of the SSVEP response using the JTFA analysis is very intuitive.

As shown in Figure 6, there is predominance in the spatial distribution of the SSVEP response but at least half of the CSP distributions differ from this standard. As shown in Tables 1 and 2, the proposed user specific adaptation significantly increases the performance for subject with an average SSVEP response amplitude and also gives an operator error proof method when compared to the manual method (Figure 7). The application of the Common Spatial Patterns method to the SSVEP paradigm has proven to be reliable and to provide an effective mapping of the relevant signal. This method provides a fast and totally automated solution to the spatial filtering problem and represents a good starting point for electrodes montage optimization: aided by the proposed features ranking index (Bhattacharyya distance), the operator should be able to exclude from the user-specific montage those channels showing a spatial weight close to zero thus minimizing user's setup time for future control sessions.

The adopted signal feature demonstrated to be effective in the discrimination of the different control tasks retaining a high discriminative power; the theoretical strength of such a feature resides in its ability to take into account an increase of energy both in the fundamental and harmonic frequencies

of the stimulation. The autoscaling property is crucial when used for providing the user with a biofeedback independent from the calibration of the classification module; in this way we propose an alternative way to solve the common *man-machine learning dilemma* (MMLD) already discussed in [25] and stating that two systems (man and machine) are strongly interdependent but have to be adapted independently in order to obtain a well-balanced mutual adaptation.

The goal of maximizing the communication performances was reached by guaranteeing an average bit-rate of approximately 51.5 bit/min with a transfer rate peak at about 70 bit/min for more than one tested subject. The overall system robustness with respect to false positives was achieved by means of an average error rate across subjects of about 2.5% even using a data analysis window of 2-second length. The three-second length analysis window has proven to be a reasonable trade-off between maximization of the information transfer-rate and minimization of the error though, with many subjects, the use of a 2-second window led to a dramatic gain in communication performances yet showing an error rate almost close to zero. Regarding this, it is also important to point out that every tested subject was facing for his/her very first time a BCI-based communication system and that each user performed the whole experimental session only once.

Maintaining that the dimensionality of the G1 and G2 control groups prevents any robust statistical analysis (which is also beyond the scope of this work); there is no evidence of performance differences between healthy and pathological subjects.

Future works will include the evaluation of the performance of the system with manually and automatically optimized electrodes montage and investigations of the

attainable control using user-specific stimulation frequency in the gamma range ( $>30$  Hz) which could theoretically lead to a less annoying and safer stimulation thanks to vision persistence.

## Acknowledgments

This work is supported by an Istituto Italiano di Tecnologia (IIT) fund. Sergio Parini's project is supported by a grant of Istituto Italiana di Tecnologia (IIT). The authors would also like to thank Engineer Gianluigi Reni and his staff, Dott.sa Elena Beretta, Paolo Perego, TBM-Lab, and WoWS!-Lab, Sensibilab, and Campus Point (LECCO, Italy) for the precious help and support.

## References

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, "Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components," *Journal of Neural Engineering*, vol. 2, no. 4, pp. 123–130, 2005.
- [3] X. Gao, D. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 137–140, 2003.
- [4] E. C. Lalor, S. P. Kelly, C. Finucane, et al., "Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 19, pp. 3156–3164, 2005.
- [5] M. Middendorff, G. McMillan, G. Calhoun, and K. S. Jones, "Brain-computer interfaces based on the steady-state visual-evoked response," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 211–214, 2000.
- [6] B. Z. Allison, D. J. McFarland, G. Schalk, S. D. Zheng, M. M. Jackson, and J. R. Wolpaw, "Towards an independent brain-computer interface using steady state visual evoked potentials," *Clinical Neurophysiology*, vol. 119, no. 2, pp. 399–408, 2008.
- [7] L. Maggi, S. Parini, P. Perego, and G. Andreoni, "BCI++: an object-oriented BCI prototyping framework," in *Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course*, Graz, Austria, September 2008.
- [8] P. Perego, L. Maggi, S. Parini, and G. Andreoni, "A home automation interface for BCI application validated with SSVEP protocol," in *Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course*, Graz, Austria, September 2008.
- [9] M. A. Pastor, J. Artieda, J. Arbizu, M. Valencia, and J. C. Masdeu, "Human cerebral activation during steady-state visual-evoked responses," *The Journal of Neuroscience*, vol. 23, no. 37, pp. 11621–11627, 2003.
- [10] C. S. Herrmann, "Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Experimental Brain Research*, vol. 137, no. 3–4, pp. 346–353, 2001.
- [11] R. S. Fisher, G. Harding, G. Erba, G. L. Barkley, and A. Wilkins, "Photic- and pattern-induced seizures: a review for the epilepsy foundation of america working group," *Epilepsia*, vol. 46, no. 9, pp. 1426–1441, 2005.
- [12] P. Martinez, H. Bakardjian, and A. Cichocki, "Fully online multicommand brain-computer interface with visual neuro-feedback using SSVEP paradigm," *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 94561, 9 pages, 2007.
- [13] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms and Applications*, Prentice-Hall, Upper Saddle River, NJ, USA, 1996.
- [14] Z. J. Koles, "The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG," *Electroencephalography and Clinical Neurophysiology*, vol. 79, no. 6, pp. 440–447, 1991.
- [15] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.
- [16] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K.-R. Müller, "Optimizing spatio-temporal filters for improving brain-computer interfacing," in *Advances in Neural Information Processing Systems*, vol. 18, pp. 315–322, MIT Press, Cambridge, Mass, USA, 2006.
- [17] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41–56, 2008.
- [18] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, New York, NY, USA, 1990.
- [19] E. Choi and C. Lee, "Feature extraction based on the Bhattacharyya distance," *Pattern Recognition*, vol. 36, no. 8, pp. 1703–1709, 2003.
- [20] C. Lee and E. Choi, "Bayes error evaluation of the Gaussian ML classifier," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1471–1475, 2000.
- [21] J. H. Friedman, "Regularized discriminant analysis," *Journal of the American Statistical Association*, vol. 84, no. 405, pp. 165–175, 1989.
- [22] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proceedings of the 13th International Conference on Machine Learning (ICML '96)*, pp. 148–156, Bari, Italy, July 1996.
- [23] L. Maggi, L. Piccini, S. Parini, G. Andreoni, and G. Panfilì, "Biosignal acquisition device—a novel topology for wearable signal acquisition devices," in *Proceedings of the 1st International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS '08)*, vol. 2, pp. 397–402, Funchal, Portugal, January 2008.
- [24] J. Kronegg, S. Voloshynovskiy, and T. Pun, "Analysis of bit-rate definitions for brain-computer interfaces," in *Proceedings of the 11th International Conference on Human-Computer Interaction (HCI '05)*, pp. 40–46, Las Vegas, Nevada, USA, June 2005.
- [25] G. Pfurtscheller and C. Neuper, "Motor imagery direct communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.