



3-D Interface for the P300 Speller BCI

Saman Noorzadeh, Bertrand Rivet, Christian Jutten

► To cite this version:

Saman Noorzadeh, Bertrand Rivet, Christian Jutten. 3-D Interface for the P300 Speller BCI. IEEE Transactions on Human-Machine Systems, 2020, 50 (6), pp.604-612. 10.1109/THMS.2020.3016079 . hal-03078380

HAL Id: hal-03078380

<https://hal.science/hal-03078380>

Submitted on 16 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

3D Interface for the P300 Speller BCI

Saman Noorzadeh, Bertrand Rivet, and Christian Jutten

Abstract—A brain-computer interface (BCI) is a common device for communication between the human brain and a computer. This paper investigated the efficiency of using a 3D interface for BCI machines. For this purpose, the P300 speller—a BCI device that enables the user to spell characters on a screen using brain waves—was modified. The classical virtual keyboard of the P300 speller was replaced with 3D stereoscopic images, which enhances the ergonomic features of the device. Moreover, the flashing paradigm on a 3D interface can affect the performance of the device in three ways: accuracy, speed and capacity. This paper proposed two different flashing paradigms called the *natural 3D* and *parallel 2D* interfaces and studied their effects regarding the three mentioned measures. The former flashes the planes in 3D space, and the latter comprises flashes of parallel keyboards at different 3D depths. The theoretical analysis of these effects is presented. The results were validated by experimental data obtained from real subjects and were compared with the classical 2D interface. Both presented keyboards increase the speed of the device, while parallel 2D has a better total performance than natural 3D.

Index Terms—brain-computer interface, P300 Speller, 3D, three-dimensional stereoscopic Images, 3D virtual keyboard

I. INTRODUCTION

BRAIN-computer interfaces (BCIs) are a direct communication link between the human brain and an external device that was initially designed to assist people with severe motor neuron diseases. Although these patients have lost their movement abilities, their cognitive abilities remain intact. Therefore, BCIs are a new technology that allows communication or the control of external devices that works based on brain activities [1], [2]. In addition, BCIs are currently considered in game navigation [3] or other virtual environments [4], serious games [5], neurofeedback [6] to manage certain disorders or devices for well-being. The P300 speller is among the BCI devices that have allowed the user to enter data (characters) with his/her brain. The user chooses a specific target symbol/character on the computer screen by concentrating on that symbol. The brain activities would be different for the attended and ignored symbols, and the computer can detect the attended symbol by processing the brain activities. The brain activity is recorded in a non-invasive manner by electroencephalogram (EEG) signals [7], and the voltage fluctuation of interest in the EEG signal owing to an evoked neural activity is called the ERP (event-related potential) [8].

The P300 speller device is based on the elicitation of ERP waves in the EEG signal that are evoked by an external stimulus. This device was first developed by Farwell and Donchin

in 1988 [9]. It works according to the oddball paradigm: if a person is subjected to an external stimulus, which is unpredictable for that person, the P300-evoked potentials are elicited in the EEG signal as a positive wave approximately 300 ms after the stimuli. The first proposed interface of the P300 speller was a 6×6 matrix of symbols with sequential, random flashes on rows and columns as external stimuli, as shown in Fig. 1. The user concentrates on one of the symbols by counting the number of flashes of that symbol and ignores other flashes. Therefore, target P300 ERPs appear in the EEG when the row and column of the target symbol are flashed. The detection of P300 waves in the recorded EEG leads to the detection of the attended symbol as the intersection of the target row and target column that caused the P300.

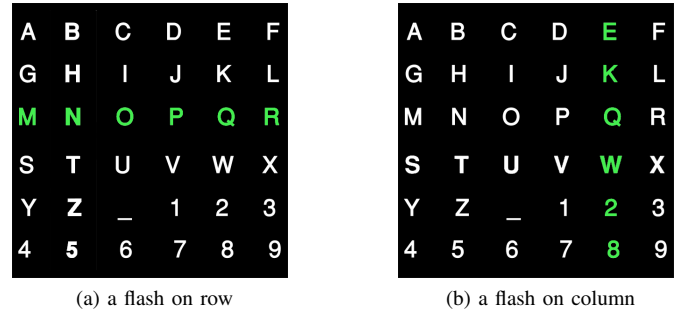


Fig. 1. The classical 2D interface with 36 symbols. If the letter 'Q' is intended to be spelled, the indicated flashes can detect this symbol.

Studies on the P300 speller can be categorized into different enhancements of the device. Some studies have proposed methods to enhance the signals recorded from subjects to remove artifacts [10]–[13], and some have focused on differentiating P300 from non-P300-evoked potentials [14], [15]. However, the BCI devices have the drawbacks of low speed and information throughput, and studies have been investigated to improve the capacity of the device. Therefore, not only the signal processing aspect but also a higher level modification can be helpful, and the interface of the device can be modified to gain a better speed or accuracy [16]. Among the recent modifications of the P300 speller interface, the flashing paradigms have been replaced by other types—e.g., rotations, translation, and zooming—making the device more comfortable for the user [17]. Similarly, others have used variations in motion, colors or flash patterns [18]–[21]. Some other studies have divided the symbol matrix to submatrices to reduce the number of flashes [22]. In other studies, the row/column flashes were replaced by overlaying faces with characters for better performance of the face paradigm than the classical flashing paradigm [23], [24]. Another paradigm, called the checkerboard paradigm was proposed [25], in which the standard matrix of symbols is virtually superimposed on a

S. Noorzadeh works at the Institute of Medical Science and Technologies at Shahid Beheshti University. She worked at GIPSA-LAB at the time of this study.

B. Rivet and C. Jutten work at GIPSA-LAB, Saint Martin d'Hères, France.

checkerboard to avoid the erroneous detection of a character and its adjacent one. For almost the same reason, [26] also proposed a new flashing paradigm, which will be used in our 3D virtual keyboards. The use of different matrix sizes has been investigated [27], and another study [28] has investigated the use of “multiple flashes” on an 8×8 grid to show that this paradigm can produce more robust ERP. More recently, the P300 speller was designed with 3D cube symbols to study the evoked P300s and compare them with the classical interface [29]. This study showed that the device outperforms the classical 2D device in accuracy and capacity while requiring a weaker workload for the user.

The present study aimed to investigate the use of a novel 3D interface in the BCI and tests a 3D setting in the P300 speller device. A 3D P300 speller can first affect the user’s acceptability of the device, a factor that is not considered in recent studies [30]. As mentioned before, BCI is not only used for patients but also in the game and leisure industry, where it has gained an essential status; therefore, for such applications, a 3D interface can be more acceptable for the user. Furthermore, the flashing paradigm on a 3D interface would be different from the classical row and column flashing. This new paradigm promotes differences in the number of flashes or trials. Therefore, we first propose 2 new paradigms and perform theoretical and experimental analyses to examine how they affect the accuracy, speed and throughput of the device. The main purpose of this study was to prove the functionality of using 3D interfaces in the P300 speller.

This paper is organized as follows: In section II, the implementation of the 3D interfaces of the P300 speller is discussed. In section III, the theoretical analysis of the proposed 3D extensions is presented. Section IV explains the collected dataset and the required signal processing and classification part. The results obtained with the experiments are then explained in section V. Finally, section VI summarizes the techniques that are presented in this study.

II. 3D EXTENSIONS OF P300-BASED BCI

Here, the interface of the P300 speller is modified by expanding the 2D keyboard to a 3D one. This modification can improve the device according to its speed, accuracy, and the information transfer rate. The generalization of the classical 2D interface to a 3D one is performed by reorganizing the symbols in a three-way table, as shown in Fig. 2. The implementation of 3D stereoscopic images (section II-1) follows this section. Afterward, two flashing strategies, called *natural 3D* (section II-A) and *parallel 2D* (section II-B), are explained in the following sections.

1) Implementation of the 3D Stereoscopic P300 speller:

The 3D device is implemented considering the stereopsis—how to represent the 3D data in two separate images for each eye [31]. For the virtual keyboard of the P300 speller, 3D stereoscopic images are implemented according to the depth cues: monocular cues [32] and binocular cues [33], [34]. The keyboard is implemented as a pair of left and right images with a specific distance between their corresponding pixels. This distance is called disparity.

There are different methods available to render stereo pairs of 3D images by positioning the cameras e.g.: *Off-axis* or *Toe-in* [35]. The toe-in method is not used owing to the discomfort it may cause for the visual system [36]. The method that is used here is the off-axis [37], [38].

Generally, viewing positive disparities are easier for the visual system than negative disparities [39]; therefore, we have implemented our virtual keyboard such that the first depth is on the screen with zero disparity, and other depths appearing on the screen have positive disparities. Other parameters, such as the distance between the two cameras, distance from the screen, and maximum disparity [40], were also determined mainly considering the user’s visual comfort [41]. An implementation of a 3D keyboard using 3 rows and 3 columns in 3 depths is shown in Fig. 2. The flashes are implemented by

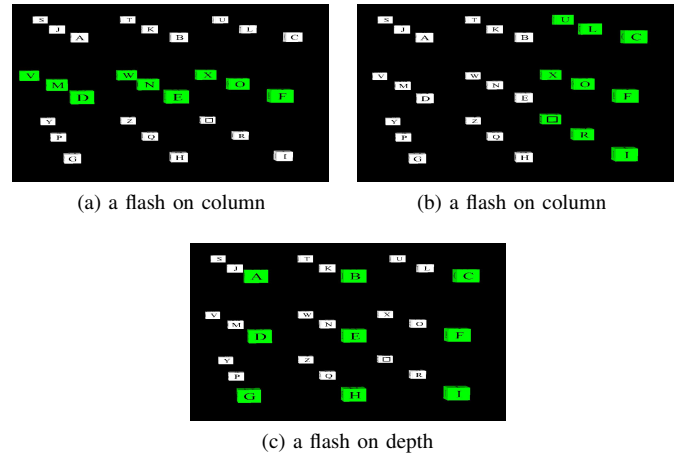


Fig. 2. Natural 3D virtual keyboard with 3 rows, 3 columns, and 3 depths. Stimulus flashes corresponding to the planes of *Natural 3D* the virtual keyboard of $3 \times 3 \times 3$ dimension. The three depicted target flashes indicate letter ‘F’.

changing the color of symbols to green and increasing the size, as shown in Fig. 2. Notably, to cover all symbols in a symbol matrix, all rows, columns, and (in the 3D case) depth planes are flashed once randomly. A repetition is the set of all these flashes. Because one repetition is not usually sufficient for a reliable detection, several repetitions are considered as a random permutation of flashes on all rows and columns.

A. Natural 3D

The natural generalization of row and column flashes of the 2D interface is the plane flashes in the 3D interface (Fig. 2), which is referred to as *natural 3D*. In such a flashing strategy, each symbol of the three-way matrix is the intersection of a row, a column, and a depth plane.

B. Parallel 2D

Another flashing approach is proposed by considering each depth of the 3D keyboard as a separate 2D keyboard. In other words, separate 2D keyboards are placed at different depth levels and function in parallel with the sequential on and off flash times. The depth is coded by a delay between the stimuli flashes of different layers, and the flashes of each layer follow

the classical row/column paradigm. Fig. 3 illustrates the onset and offsets of the flashes of an interface with two layers by showing the stimuli timelines: the dotted line corresponds to one of the depth levels, and the continuous line corresponds to the other, where, for N stimuli, t_i , $i \in [1, N]$, is the onset of the i^{th} flash on the first depth layer and t'_i is the onset of the i^{th} flash on the second depth layer. The interlacing flashes lead to independent or parallel layers. This new configuration of interface and flashes is called *parallel 2D*.

The number of 2D keyboards that can function in parallel is limited because of the limited frame rate of the imaging device and to ensure good detection; the delay between layers should not be too small.

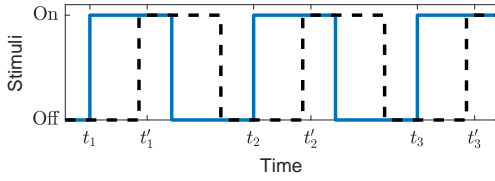


Fig. 3. Parallel flashes on 2 depth levels in the parallel 2D interface. The solid and dashed lines represent the times of the stimuli for the first and second depths, respectively.

To address the small delays between flashes of depths in this interface, a different paradigm is applied here: Among any pair of repetitions, any two consecutive flashes should be different. Thus, let $f_k^{(j)}(i)$ denote the row or column index of the i -th flash of the j -th repetition on the k -th layer. Then, with a two-depth interface ($k = 1$ or 2), and any $i, i' \in [1, N]$ and any j, j' repetition, the sequence of flashes on the first and second layers must verify that:

$$\begin{aligned} \left(f_1^{(j)}(i), f_2^{(j)}(i) \right) &= \left(f_1^{(j')}(i'), f_2^{(j')}(i') \right) \\ \Rightarrow (i, j) &= (i', j'). \end{aligned} \quad (1)$$

This equation expresses that the same pair of consecutive flashes cannot occur in two different repetitions. Because the classifier score will be computed cumulatively among repetitions, by having the mentioned paradigm of eq. (1), the false detected target flash will not gain more scores through the repetitions. This implementation is necessary because the delay between successive flashes is small. Additionally, the keyboards in different depths only work in parallel and do not increase the ISI (inter-stimuli interval) to keep the speed intact. However, if such a paradigm had not been implemented, the target flash would likely have been detected erroneously as its previous or next flash.

Another flashing paradigm is also used on row/column flashes to avoid erroneous detection of adjacent symbols. This is fully presented in [26].

III. THEORETICAL ANALYSIS OF THE 3D EXTENSIONS

Having proposed the 3D extensions of the virtual keyboard of the P300 speller and flashing paradigms, we will provide the theoretical analysis concerning the changes in accuracy, speed, and capacity.

A. Accuracy

As mentioned in the Introduction section, the classical P300 speller has a 2D interface (see Fig. 1), which we consider as the reference to compare with the new 3D extensions. First, we review the classical 2D interface for its accuracy. The target symbol is the intersection of a target row and a target column. Let's define R and C as discrete random variables representing the detected rows and columns, respectively. These random variables have a uniform distribution because the flashes on the rows and columns are equal and we expect equal P300 elicitation. We assume that the detection of a row and a column is independent but equal to $P \in [0, 1]$. Now, if $p(R = r, C = c)$ denotes the probability that the row r and column c are jointly detected, equation (2) shows the probability that the target symbol is correctly detected, indicating that the target row, r_t , and target column, c_t , are both detected correctly.

$$\begin{aligned} Acc_{classical2D} &= p(R = r_t, C = c_t) \\ &= p(R = r_t)p(C = c_t) = P^2. \end{aligned} \quad (2)$$

In the natural 3D interface, the probability of detecting the symbol at the intersection of a row plane r , column plane c , and depth plane d is expressed as $p(R = r, C = c, D = d)$, where R , C , and D are uniform random variables. We assume equal numbers of rows, columns and depths, equal to the number of rows (columns) in classical 2D. Additionally, assuming that the marginal detections are independent (i.e., $p(R = r, C = c, D = d) = p(R = r)p(C = c)p(D = d)$) and equal to P , the accuracy of the natural 3D interface is given by:

$$Acc_{natural3D} = p(R = r_t, C = c_t, D = d_t) = P^3, \quad (3)$$

where r_t , c_t , and d_t are the target row plane, column plane, and depth plane, respectively. The natural 3D accuracy (eq. (3)) is theoretically lower than the classical 2D accuracy (eq. (2)).

Finally, for the parallel 2D approach, as noted in section II-B, the probability of detecting a symbol as the intersection of row r and column c in layer d is independent of d : $p(R = r, C = c, D = d) = p(R = r, C = c)$. This is because the keyboards are working in parallel. With factorizing probabilities, the accuracy of the parallel 2D interface is given by:

$$Acc_{parallel2D} = p(R = r_t, C = c_t, D = d_t) = P^2. \quad (4)$$

Considering $P \in [0, 1]$ as the marginal probability, the accuracy of the classical 2D and parallel 2D was shown to be equal to P^2 ; however, that of the natural 3D interface would be lower and equal to P^3 for the same number of rows and columns (and depth for 3D).

B. Speed: number of flashes

Among the most important characteristics of the newly proposed interfaces is their speed. Each P300 speller interface requires, in each repetition, a minimum number of flashes. The number of these flashes estimates the speed of the interface. Here, we compare the theoretical speed of the interfaces by calculating the number of required flashes.

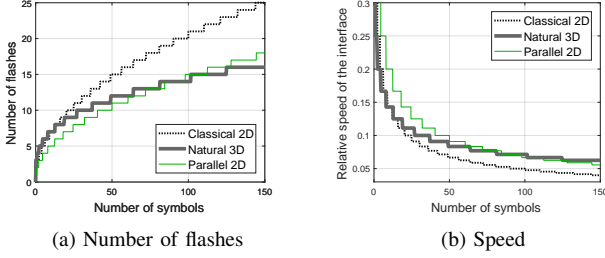


Fig. 4. Comparison of the speed of the interfaces. Fig. 4(a) shows the minimum number of flashes—as calculated in equations (5),(6), and (7)—against the total number of symbols on the interface. The number of depths in the parallel 2D interface is 2 ($d = 2$). Fig. 4(b) shows the relative speed of the interfaces defined in eq.(8), when $T = 1$.

With a total number of S symbols on the interface of the classical 2D, we can have the minimum number of flashes if the symbols are placed in a square-shaped matrix; therefore, there are $S^{\frac{1}{2}}$ number of rows and also the same number of columns, leading to $2 \times S^{\frac{1}{2}}$ number of rows and column flashes. The minimum number of flashes required to cover all the columns and rows in one repetition is hence calculated from:

$$N_{2D} = \left\lceil 2 \times S^{\frac{1}{2}} \right\rceil, \quad (5)$$

where $\lceil \cdot \rceil$ is the ceiling function. In the row-column-depth configuration for natural 3D, the minimum number of flashes is calculated from eq. (6), where $S^{\frac{1}{3}}$ is referred to as the number of rows, or columns, or depths in a cubic matrix of symbols.

$$N_{natural3D} = \left\lceil 3 \times S^{\frac{1}{3}} \right\rceil. \quad (6)$$

For parallel 2D, the total number of symbols can be distributed over the number of layers, referred to as d , so that each layer contains $\frac{S}{d}$ symbols. The minimum number of flashes is then the number of row/column flashes on each layer of the interface, and is calculated as:

$$N_{parallel2D} = \left\lceil 2 \times \left(\frac{S}{d} \right)^{\frac{1}{2}} \right\rceil, \quad (7)$$

Fig. 4 compares the minimum number of flashes on each interface considering 2 layers for parallel 2D ($d = 2$). With more than nine symbols in the interface ($S > 9$), the classical 2D interface requires more flashes than the natural 3D interface; when there are fewer than 85 symbols, the parallel 2D interface needs fewer flashes than the natural 3D interface. However, the parallel 2D always requires fewer flashes than the classical 2D one.

As noted above, the speed of the interface is directly related to the number of flashes, N . By defining the speed F of the interface inversely proportional to the duration of each repetition, its expression is given by:

$$F = \frac{1}{N \times T}, \quad (8)$$

where T is the inter-stimuli interval (ISI) in seconds. Thus, the smaller is number of flashes, the faster is the interface (Fig. 4(b)).

C. Capacity

A more relevant performance measure used to quantize the information transfer rate (ITR) is the capacity. According to [42], for one repetition, the capacity of a S -symbol keyboard, with an accuracy of Acc , is defined as:

$$Q = \log_2(S) + Acc \log_2(Acc) + (1 - Acc) \log_2\left(\frac{1 - Acc}{S - 1}\right), \quad (9)$$

The capacity Q is expressed in bits per repetition. However, the duration of repetitions can be different for different interfaces because of the different number of flashes that each interface might have. Thus, the comparison between interfaces is more appropriate if we consider the time and bit-rate B , in bits per second. This is defined as:

$$B = Q \times F = \frac{Q}{N \times T}, \quad (10)$$

or the relative bit-rate B' , expressed in bits per flash because the ISI is the same for all flashes:

$$B' = \frac{Q}{N}. \quad (11)$$

As shown in Figs. 5(d), 5(e), and 5(f), the bit-rates of the natural 3D and parallel 2D interfaces are always larger than that of the classical 2D interface for a given accuracy. However, the bit-rate of the natural 3D interface is only larger than that of the classical 2D for large values of marginal probability P (Figs. 5(a), 5(b), and 5(c)). This threshold value for P decreases as the number of symbols increases because the relationships between the accuracy and marginal probability of good detection is nonlinear.

However, the bit-rate of the parallel 2D interface is always larger than that of the classical 2D interface because of the smaller number of flashes with the same accuracy. This improvement factor is given by $N_{2D}/N_{parallel2D}$.

IV. EXPERIMENTS AND PROCESSING METHODS

Our primary goal is the possibility of designing a 3D interface for the P300 speller and designing the flashing paradigms for this new interface. Considering the previous points, experiments on subjects were performed to analyze the theoretical analysis using real data. Because we aimed to prove the functionality of a 3D setting in a BCI device, the experiments presented here are preliminary with a small number of subjects. We mainly intended to link what we concluded in theory to the results observed in real subjects. Based on this study, future studies can include more experiments in real-time and/or with a larger number of subjects.

A. Dataset

The data were collected from 16 volunteers between 22 and 34 years of age (with a mean age of 27 years), with normal stereoscopic vision and a normal neurological state. Eight participants among the 16 were subjected to the natural 3D keyboard, and the other eight were subjected to the parallel 2D keyboard. Both sets of subjects also were subjected to the

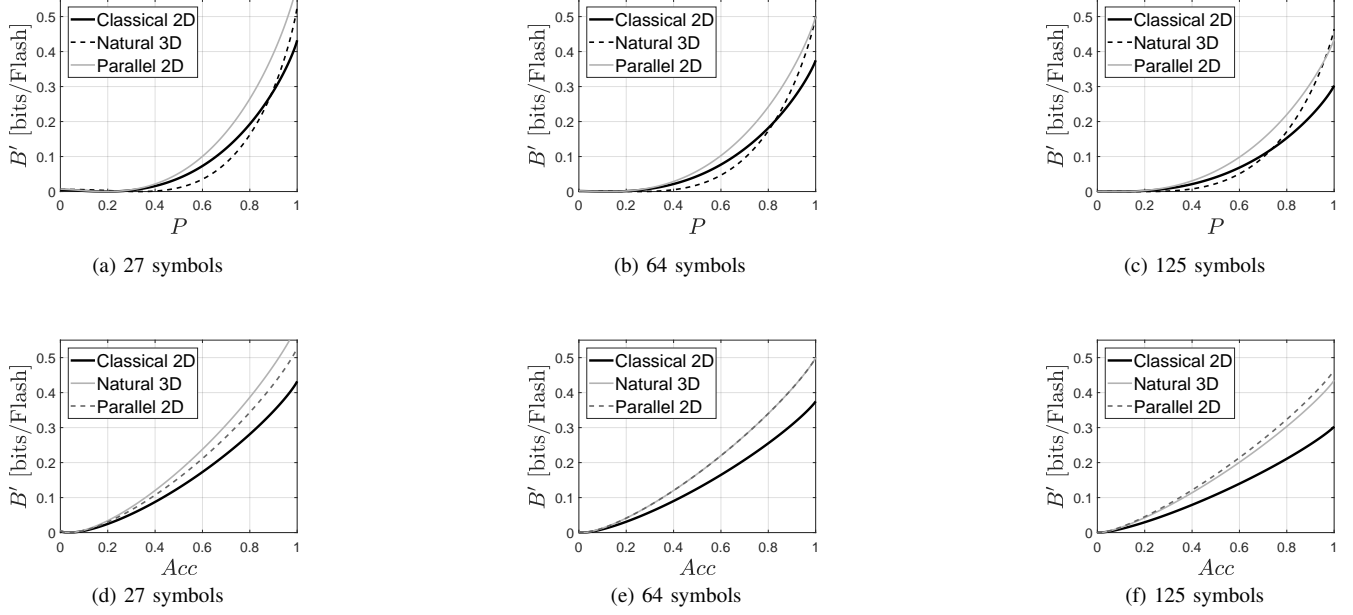


Fig. 5. Relative bit-rate per flash B' of eq. (11) for several symbols S in the interface against the marginal probability of good detection P (top Figs.) and against the accuracy (bottom Figs.). In Fig. 5(e), $B'_{natural3D} = B'_{parallel2D}$ because $N_{natural3D} = N_{parallel2D}$ for 64 symbols.

classical 2D interface, so that the results can be compared with the 3D ones.

Each subject recorded two sessions for each interface. In the first session, he/she had 50 predefined symbols to type on the keyboard. In the second session, the subjects were free to choose a word to spell on the interface. All the analyses were performed offline.

The interfaces were implemented using OpenGL. The number of symbols on the interfaces was almost the same to make a fair comparison; therefore, the classical 2D interface is a 6×6 matrix of symbols, the natural 3D keyboard is in 3 rows, 3 columns, and 3 depths with 27 symbols, and the parallel 2D keyboard has 32-character placements in 2 depths of 4×4 matrices of symbols. The number of rows and columns was chosen to have almost the same number of symbols on different interfaces. The number of symbols cannot be exactly the same for all keyboards because the number of symbols versus the number of rows, columns and depth is a stepwise function (Fig. 4). In other words, the dimension of the interface matrix—which is already defined to be cubic—defines the number of positions that the symbols can take and can be detected with the P300 speller.

The ISI is equal to 133 milliseconds with a duty cycle of 50% for the flashes. To spell each letter, 5 repetitions are considered. The subjects sit at a distance of 45 cm from a 3D television.

The EEG from the subjects was recorded via 16 active electrodes using a g.USBamp device from g.tec [43]. The signal is sampled at a rate of 1200 Hz.

B. Signal Processing

The EEG signals recorded during the experiments are first filtered by a bandpass filter in the frequency band between

1 and 12 Hz using a fourth-order Butterworth filter. Then, to enhance the signals, spatial filters are estimated [12], [44] from the training data. This filter is applied to enhance the P300-evoked potentials. The spatial filter used here is xdown, which is fully described in [45].

C. Classification

The temporally and spatially filtered signals are then categorized into two parts according to the epoch type, which can be target or non-target, and is used as the two-class data to train the classifier. As mentioned, using spatial filters, the data are already projected onto the ERP subspace to enhance the ERP signals and improve the signal-to-noise ratio [45]. Therefore, the spatially filtered target and non-target ERPs are directly used as the features. In this study, Bayesian Linear Discriminant Analysis (BLDA) was used [46].

In the test phase, the trained spatial filters are applied on the signal, and the predictive score for each stimulus flash is provided by the BLDA classifier. Each row and column (and depth in natural 3D) will receive a score in each repetition. These scores are summed to predict the symbol. Let $S_k(i, j)$ be the predictive probability score ($0 \leq S \leq 1$) of the i -th row (or column or depth) flash of the j -th repetition on the row, column or depth, $k \in [r, c \text{ or } d]$.

The spelled symbol is then predicted according to the intersection of a row and a column (and the depth for the natural 3D), which have the largest scores summed over all the repetitions:

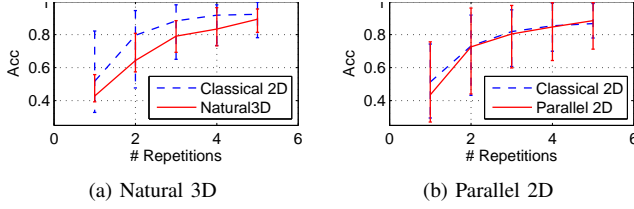


Fig. 6. Accuracy against the number of repetitions for *Natural 3D* (Fig. 6(a)) and *Parallel 2D* (Fig. 6(b)) interfaces. The median accuracies are plotted, and the errorbars extend from the 10% quantile to the 90% quantile.

$$\hat{k} = \underset{i}{\operatorname{argmax}} \sum_{j=1}^J S_k(i, j), \quad k \in [r, c \text{ or } d], \quad (12)$$

$$\text{symbol} = \text{intersect}(\hat{r}, \hat{c}, \hat{d}). \quad (13)$$

Because the number of subjects is small, to have more accurate results, we have two strategies for validation. First, a K-fold cross-validation is adopted using the first-session recordings. As already mentioned, this session comprises 50 pre-defined symbols, so the database is divided into 40 symbols for training and 10 symbols for the tests. This partitioning is performed randomly 2,000 times, and the results are averaged. Section V demonstrates the results of such cross-validation. These results are then summarized in Table II.

The next strategy is the train-test in which all 50 symbols are used for training, and the second session recording is used for the test. To calculate the accuracy, the subject is asked what word he/she has spelled after the experiment. These results are summarized in section V Table III

V. RESULTS AND DISCUSSIONS

The experimental results offer two comparisons that are presented in this section: the classification accuracy and capacity of the presented interfaces. The speed of the interface depends on the implementation of the 3D interfaces and is directly measured by the ISI and number of repetitions. Thus, without the experiments, the speed of different interfaces has already been compared theoretically in the previous sections.

A. Accuracy and marginal probability

Using the data from the experiments, the classification accuracies of the two 3D extensions are compared in Fig. 6. It should be pointed out that the classification accuracy increases with the number of repetitions.

The classification accuracy of the classical 2D interface is larger than that of the natural 3D interface (see also Table I: the first and second rows for t-test p-values). The accuracy of classical and parallel 2D interfaces are more similar (Table I: row 3). Parallel 2D accuracy also outperformed the classical 2D approach, as expected, in the theoretical part.

Because the obtained results do not match completely with the theoretical part (especially for natural 3D), we must evaluate the assumption of the equality of marginal probabilities. To verify the theoretical assumptions on the marginal

probabilities, P , of the interfaces (section III-A), the accuracy of detecting rows and columns (and depths) is separately calculated in the experiments. The results are then depicted in Fig. 7, which shows that the equal marginal probability hypothesis cannot be completely achieved for all dimensions with the natural 3D interface. In particular, based on the hypothesis tests of Table I: row 4 and row 5, the marginal probability of good detection of the depth is smaller than the marginal probabilities of good detection of the rows or columns (Table I: row 5). By contrast, the two marginal probabilities of good detection of the rows or columns are the same (Table I: row 6) and are almost similar to those of good detection using the classical 2D interface (Table I: row 7 and 8). These results were more comparable in the last repetitions. One can interpret the lower accuracy of depth detection than those of rows and columns because the non-target flashed depth layers in different visual fields may interrupt the subject while focusing on the intended symbol. Another advantage of the 3D interface is that, in the very first repetitions, the detection of rows and columns is more accurate than that using classical 2D because of the lower dimension of the interface and higher amplitudes of P300 waves in the 3D stereoscopic interface [29].

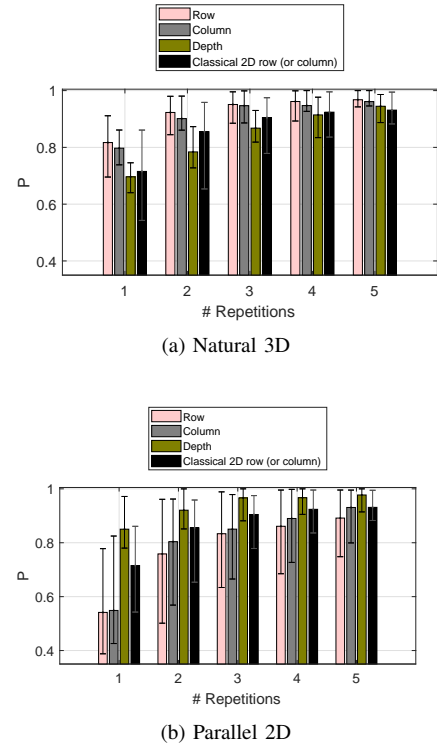


Fig. 7. Marginal probability (P) against the number of repetitions for *Natural 3D* and *Parallel 2D* interfaces. The median accuracies are plotted, and the error bars extend from the 10% quantile to the 90% quantile.

On the parallel 2D interface, the two assumptions regarding the marginal probabilities (identical and independent marginal probabilities on the rows and columns, and a marginal probability of detecting the depth close to one; see section III-A) are well verified. Both distributions of marginal probabilities on the rows and columns are similar in mean (Table I:row

Row no.	Null hypothesis	paired t-test type	P-values for 5 No. of repetitions				
			1	2	3	4	5
1	$\text{acc}(\text{C2D}) = \text{acc}(\text{N3D})$	both-tailed	0.05	0.03	.05	0.07	0.10
2	$\text{acc}(\text{C2D}) > \text{acc}(\text{N3D})$	right-tailed	0.98	0.99	0.98	0.97	0.95
3	$\text{acc}(\text{C2D}) = \text{acc}(\text{P2D})$	both-tailed	0.77	0.66	0.84	0.52	0.74
4	$P(\text{depth}(\text{N3D})) < P(\text{row}(\text{N3D}))$	left-tailed	0.99	1.00	1.00	1.00	0.99
5	$P(\text{depth}(\text{N3D})) < P(\text{column}(\text{N3D}))$	left-tailed	1.00	1.00	0.99	0.96	0.97
6	$P(\text{row}(\text{N3D})) = P(\text{column}(\text{N3D}))$	both-tailed	0.70	0.88	0.74	0.66	0.93
7	$P(\text{row}(\text{C2D})) = P(\text{row}(\text{N3D}))$	both-tailed	0.47	0.90	0.85	0.61	0.82
8	$P(\text{column}(\text{C2D})) = P(\text{column}(\text{N3D}))$	both-tailed	0.96	0.60	0.56	0.61	0.81
9	$P(\text{column}(\text{P2D})) = P(\text{row}(\text{P2D}))$	both-tailed	0.08	0.06	0.36	0.14	0.15
10	$P(\text{row}(\text{C2D})) = P(\text{row}(\text{P2D}))$	both-tailed	0.07	0.41	0.32	0.22	0.32
11	$P(\text{column}(\text{C2D})) = P(\text{column}(\text{P2D}))$	both-tailed	0.19	0.64	0.35	0.33	0.53
12	$P(\text{depth}(\text{C2D})) = 1$	left-tailed	0.001	0.006	0.014	0.010	0.012

TABLE I

Paired t-test p-values to compare the accuracy of the devices. THE TABLE ROW NUMBERS SHOW THE ACCURACIES, WHICH ARE COMPARED. THIS INFORMATION IS FOUND IN THE TEXT. COMPARING THE TWO SETS OF DATA, X AND Y, FOR BOTH TAILED T-TESTS, THE NULL HYPOTHESIS IS THAT DATA IN $x - y$ ARE SAMPLES OF A NORMAL DISTRIBUTION WITH A MEAN OF 0; IN THE RIGHT-TAILED T-TEST WITH MEAN GREATER THAN 0, AND IN THE LEFT-TAILED TEST WITH A MEAN LESS THAN 0.

9)(Fig. 7(b)), and they are also similar to the marginal probability of the classical 2D interface (Table I:row 10 for rows and Table I:row 11 for columns). Moreover, the marginal probability of detecting the depth in parallel 2D is close to one after at least 3 repetitions (Table I:row 12). This analysis validates the assumption that the depth can be efficiently coded by the time shifts between the layers and orthogonal sequences of the flashes (eq. (4)). For information about the statistical tests, refer to [47].

Lastly, the cross-validation and the train-test results are summarized in tables II and III.

Interface	Accuracy (%)	Capacity (bit-rate per flash)
Classical 2D	79.5 ± 4.04	70.03 ± 6.75
Natural 3D	62.6 ± 5.4	47.88 ± 7.90
Parallel 2D	77.63 ± 4.67	67.42 ± 8.09

TABLE II

CROSS-VALIDATION RESULTS: SUMMARY OF THE PRESENTED RESULTS ON ALL 5 REPETITIONS WITH THE CROSS-VALIDATION.

Interface	Accuracy (%)	Capacity (bit-rate per flash)
Classical 2D	75.16 ± 2.21	62.4 ± 3.33
Natural 3D	73.25 ± 3.14	58.35 ± 4.73
Parallel 2D	74.65 ± 2.2	62.25 ± 3.25

TABLE III

TEST RESULTS: SUMMARY OF ALL REPETITIONS AND SUBJECTS, WHEN A TEST SESSION WAS CONSIDERED FOR THE SUBJECTS. THE TRAIN DATA ARE USED FOR TRAINING, AND A SEPARATE TEST SESSION IS CONSIDERED FOR EACH SUBJECT (AS EXPLAINED IN SECTION IV), IN WHICH HE/SHE SPELLED A WORD OF FOUR TO EIGHT CHARACTERS FREELY.

B. Capacity

Here, using the experimental results, the capacities of the interfaces are compared against time. The capacities are depicted according to elapsed time in Fig. 8. First, having the same number of repetitions in both 3D approaches takes less time than in the classical 2D approach, as expected in the theoretical part. Second, Fig. 8(a) shows that the bit-rate of the natural 3D interface is lower than that of the classical 2D interface, likely owing to the lower accuracy of the natural 3D interface. Compared with the theoretical analysis, the accuracy of the natural 3D interface is only larger than that of the

classical 2D interface for very large marginal probabilities (almost 90% for 27 symbols), unlike the case here (Figs. 5(a) to 5(c)). Other considerations should be applied to improve the marginal probability in future work. In Figs. 5(d) to 5(f), the capacity of natural 3D is higher than that of classical 2D only if they have the same accuracy.

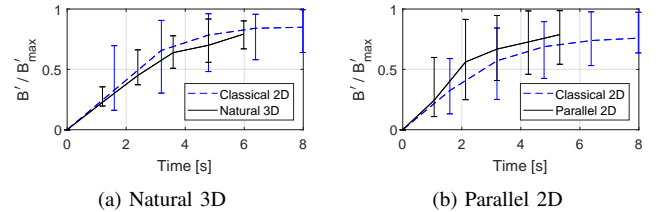


Fig. 8. Comparison of relative bit-rates. B'_{Max} is the maximum bit-rate assuming a perfect accuracy (i.e. $\text{Acc} = 100\%$). The median accuracies are plotted, and the error bars extend from the 10% quantile to the 90% quantile.

The bit-rate of the parallel 2D is higher than that of the classical 2D interface (Fig. 8 on the right). This observation was confirmed by the theoretical analysis (Fig. 5) because, for the same accuracy, the capacity of the parallel 2D was higher because of its higher speed. Finally, it should be noted that the high variance of the capacities is due to the sensitivity of the capacity to the accuracy and is averaged over different subjects; for each, there is a probability of visual or brain distraction while spelling on the interfaces.

VI. CONCLUSIONS

Two 3D extensions of the classical 2D P300 speller have been proposed in this paper: *natural 3D* and *parallel 2D*. The major aim of the extensions was to improve the efficiency of the device in 3D settings and make the P300 speller device more ergonomic for users because the ergonomics of this device has not been considered much previously. Experiments on a set of 16 subjects showed that the two new 3D approaches have some advantages compared with the classical 2D interface. One advantage is the comfort in spelling on the interface and the ergonomics. Considering these factors, each subject declared his/her comfort in a survey and the obtained results showed that the 3D keyboard can increase the user's

satisfaction with the device. Through this survey, we found that 75% of the subjects preferred the 3D keyboard to the 2D for its comfort and higher ergonomic features. Other advantages are related to the accuracy, capacity, and speed. Therefore, in this study, the performance, in terms of these 3 factors, has been theoretically discussed, and then the experiments on real data were performed to prove the theoretical results. Although the experiments involved a small number of subjects, they showed that the 3D settings can be applied in BCI applications.

Regarding the **speed**, this study showed that:

- Both 3D approaches are always faster than the classical 2D approach with the same number of symbols.
- The speed of the natural 3D interface is higher than that of the parallel 2D interface only for more than 61 characters on the interface.

Regarding the **Accuracy**, this study shows that:

- The accuracy of the natural 3D interface is lower than that of the classical 2D interface.
- The accuracy of the parallel 2D interface is almost the same as that of the classical 2D interface.
- The relationship between the accuracy of the devices ($Acc_{2D} = Acc_{parallel2D} > Acc_{Natural}$) was validated in both the theoretical and experimental results; however, the marginal probability of the depth of the natural 3D interface was shown to be less accurate than that of the other two layers in the experimental results, despite our hypothesis in the theoretical part. Thus the parallel 2D interface is recommended when considering the 3D interface.

Considering the **capacity (bit-rate)** of the different interfaces theoretically and experimentally, we can conclude that:

- The capacity of the parallel 2D interface is higher than that of the other interfaces.
- The natural 3D interface has a lower capacity than the classical 2D interface.

This study has proven the functionality of 3D interfaces in the P300 speller. According to the higher speed and accuracy, we recommend the parallel 2D interface. This finding is validated by both theoretical and experimental results. The use of any of these paradigms depends on the application in which it is being used. It should be noted that the general effect of 3D stereoscopic images on the elicitation of ERPs was not investigated in this study and can be further analyzed in a future study. Here, only the efficiency of 3D was studied in a P300 speller paradigm. Notably, the 3D settings can be used not only in the P300 devices but also in other technologies that require ergonomics, 3D environments, or even a higher speed, such as virtual worlds [48].

REFERENCES

- [1] L. R. Hochberg, M. D. Serruya, G. M. Friehe, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue, "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature*, vol. 442, no. 7099, pp. 164–171, 2006.
- [2] P. R. Kennedy and R. A. Bakay, "Restoration of neural output from a paralyzed patient by a direct brain connection," *Neuroreport*, vol. 9, pp. 1707–1711, 1998.
- [3] R. Leeb, M. Lancelle, V. Kaiser, D. W. Fellner, and G. Pfurtscheller, "Thinking penguin: multimodal brain-computer interface control of a vr game," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 5, no. 2, pp. 117–128, 2013.
- [4] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, "Brain-computer communication: motivation, aim, and impact of exploring a virtual apartment," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 4, pp. 473–482, 2007.
- [5] Q. Wang, O. Sourina, and M. K. Nguyen, "Eeg-based" serious" games design for medical applications," in *2010 International conference on Cyberworlds*. IEEE, 2010, pp. 270–276.
- [6] V. A. Grin-Yatsenko, V. A. Ponomarev, O. Kara, B. Wandernoth, M. Gregory, V. A. Ilyukhina, and J. D. Kropotov, "Effect of infra-low frequency neurofeedback on infra-slow eeg fluctuations," in *Biofeedback*. IntechOpen, 2018.
- [7] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, p. 767–791, 2002.
- [8] S. J. Luck, *An introduction to the event-related potential technique*. MIT press, 2014.
- [9] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, December 1988.
- [10] H. Serby, E. Yom-Tov, and G. Inbar, "An improved P300-based brain-computer interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 1, pp. 89–98, March 2005.
- [11] M. Kaper, P. Meinicke, U. Grossekhoefer, T. Lingner, and H. Ritter, "BCI competition 2003-data set IIB: support vector machines for the P300 speller paradigm," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1073–1076, June 2004.
- [12] B. Rivet and A. Souloumiac, "Optimal linear spatial filters for evoked potentials: an asymptotic performance analysis," *Signal Processing*, vol. 93, pp. 387–398, 2013.
- [13] B. Somers, T. Francart, and A. Bertrand, "A generic eeg artifact removal algorithm based on the multi-channel wiener filter," *Journal of neural engineering*, vol. 15, no. 3, p. 036007, 2018.
- [14] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayouduh, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A comparison of classification techniques for the P300 Speller," *Journal of Neural Engineering*, vol. 3, no. 4, pp. 299–305, December 2006.
- [15] J. T. Philip and S. T. George, "Visual p300 mind-speller brain-computer interfaces: A walk through the recent developments with special focus on classification algorithms," *Clinical EEG and neuroscience*, p. 1550059419842753, 2019.
- [16] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Saboor, and I. Volosyak, "Brain-computer interface spellers: A review," *Brain sciences*, vol. 8, no. 4, p. 57, 2018.
- [17] Y. Liu, Z. Zhou, and D. Hu, "Comparison of stimulus types in visual p300 speller of brain-computer interfaces," in *9th IEEE International Conference on Cognitive Informatics (ICCI'10)*. IEEE, 2010, pp. 273–279.
- [18] B. Hong, F. Guo, T. Liu, X. Gao, and S. Gao, "N200-speller using motion-onset visual response," *Clinical neurophysiology*, vol. 120, pp. 1658–1666, 2009.
- [19] K. Takano, T. Komatsu, N. Hata, Y. Nakajima, , and K. Kansaku, "Visual stimuli for the P300 brain-computer interface: A comparison of white/gray and green/blue flicker matrices," *Clinical neurophysiology*, 2009.
- [20] S. M. and S. F., "Visual modifications on the p300 speller bci paradigm," *J Neural Eng.*, vol. 6, p. 046011, 2009.
- [21] S. M. M. Martens, N. J. Hill, J. Farquhar, and B. Scholkopf, "Overlap and refractory effects in a brain-computer interface speller based on the visual P300 event-related potential," *Journal of Neural Engineering*, vol. 6, no. 2, p. 026003, 2009.
- [22] J.-h. Shi, J.-z. Shen, Y. Ji, and F.-l. Du, "A submatrix-based p300 brain-computer interface stimulus presentation paradigm," *Journal of Zhejiang University SCIENCE C*, vol. 13, no. 6, pp. 452–459, 2012.
- [23] Q. Li, S. Liu, J. Li, and O. Bai, "Use of a green familiar faces paradigm improves p300-speller brain-computer interface performance," *PloS one*, vol. 10, no. 6, p. e0130325, 2015.
- [24] J. Jin, B. Z. Allison, T. Kaufmann, A. Kübler, Y. Zhang, X. Wang, and A. Cichocki, "The changing face of P300 BCIs: A comparison of stimulus changes in a P300 bci involving faces, emotion, and movement," *PloS one*, vol. 7, no. 11, p. e49688, 2012.

- [25] G. Townsend, B. K. LaPallo, C. B. Boulay, D. J. Krusienski, G. E. Frye, C. K. Hauser, N. E. Schwartz, T. M. Vaughan, J. R. Wolpaw, and E. W. Sellers, "A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns," *Clinical neurophysiology*, vol. 121, no. 7, pp. 1109–1120, 2010.
- [26] H. Cecotti and B. Rivet, "One step beyond rows and columns flashes in the p300 speller: a theoretical description," *International Journal of bioelectromagnetism*, vol. 13, no. 1, pp. 39–41, 2010.
- [27] B. Z. Allison and J. A. Pineda, "Erps evoked by different matrix sizes: implications for a brain computer interface (bci) system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 110–113, 2003.
- [28] —, "Effects of SOA and flash pattern manipulations on erps, performance, and preference: implications for a bci system," *International journal of psychophysiology*, vol. 59, no. 2, pp. 127–140, 2006.
- [29] J. Qu, F. Wang, Z. Xia, T. Yu, J. Xiao, Z. Yu, Z. Gu, and Y. Li, "A novel three-dimensional p300 speller based on stereo visual stimuli," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 4, pp. 392–399, 2018.
- [30] S. Noorzadeh, B. Rivet, and C. Jutten, "Beyond 2d for brain-computer interfaces: Two 3d extensions of the p300-speller," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 5899–5903.
- [31] I. P. Howard and B. J. Rogers, *Binocular Vision and Stereopsis*. New York: Oxford Univ. Press, 1995.
- [32] F. Devernay and P. Beardsley. Springer-Verlag, 2010, ch. Stereoscopic cinema. Image and Geometry Processing for 3-D Cinematography, pp. 5–51.
- [33] T. Shibata, T. Kawai, K. Ohta, M. Otsuki, N. Miyake, Y. Yoshihara, and T. Iwasaki, "Stereoscopic 3-D display with optical correction for the reduction of the discrepancy between accommodation and convergence," *Journal of the Society for Information Display*, vol. 13, p. 665–671, 2005.
- [34] J. M. Rolfe and K. J. Staples, *Flight Simulation*. Cambridge University Press, 1986.
- [35] "Copy of stereo rendering- stereo rendering— daniel wexler, gelato developer," June 2006, http://www.nvidia.com/object/IO_36545.html.
- [36] "Stereographics theory," <http://paulbourke.net/exhibition/vpac/theory.html>.
- [37] C. Fehn, "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV," in *Proceeding of Stereoscopic Displays and Applications*, 2002.
- [38] C. Wang and A. A. Sawchukg, "Disparity manipulation for stereo images and video," in *Proceedings of Stereoscopic Displays and Applications XIX*, 2008.
- [39] M. T. Lambooi, W. A. IJsselstein, and I. Heynderickx, "Visual discomfort in stereoscopic displays: a review," in *Electronic Imaging 2007*. International Society for Optics and Photonics, 2007, pp. 64 900I–64 900I.
- [40] R. Vlad., "State-of-the-art report on stereoscopic 3D issues with focus on quality," Grenoble, 2011, deliverable MOOV3D project.
- [41] K. Ukai, "Human factors for stereoscopic images," 2006, pp. 1697–1700.
- [42] J. Kronegg, S. Voloshynovskiy, and T. Pun, "Analysis of bit-rate definitions for brain-computer interfaces," in *Proceedings of the 2005 international conference on human-computer interaction*, 2005.
- [43] "<http://www.gtec.at/products/hardware-and-accessories/g.usbamp-specs-features>."
- [44] B. Rivet, A. Souloumiac, G. Gibert, and V. Attina, "P300 speller brain-computer interface: Enhancement of P300 evoked potential by spatial filters," in *16th European Signal Processing Conference (EUSIPCO)*, 2008.
- [45] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert, "xDAWN algorithm to enhance evoked potentials: application to brain-computer interface," *IEEE Transactions on Biomedical Engineering*, vol. 56, 2009.
- [46] U. Hoffmann, J. M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects," *Journal of Neurosci Methods*, vol. 167, no. 1, pp. 115–125, 2008.
- [47] M. Billinger, I. Daly, V. Kaiser, J. Jin, B. Z. Allison, G. R. Müller-Putz, and C. Brunner, "Is it significant? guidelines for reporting bci performance," in *Towards Practical Brain-Computer Interfaces*. Springer, 2012, pp. 333–354.
- [48] J. Bayliss, "Use of the evoked potential P3 component for control in a virtual apartment," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 113–116, 2008.