35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

DTU BCI Speller: An SSVEP-based Spelling System with Dictionary Support

Adnan Vilic[1], Troels W. Kjaer[2], Carsten E. Thomsen[3], S. Puthusserypady[1], and Helge B.D. Sorensen[1]

Abstract— **In this paper, a new brain computer interface (BCI) speller, named DTU BCI speller, is introduced. It is based on the steady-state visual evoked potential (SSVEP) and features dictionary support. The system focuses on simplicity and user friendliness by using a single electrode for the signal acquisition and displays stimuli on a liquid crystal display (LCD). Nine healthy subjects participated in writing full sentences after a five minutes introduction to the system, and obtained an information transfer rate (ITR) of 21.94 ± 15.63 bits/min. The average amount of characters written per minute (CPM) is 4.90 ± 3.84 with a best case of 8.74 CPM. All subjects reported systematically on different user friendliness measures, and the overall results indicated the potentials of the DTU BCI Speller system. For subjects with high classification accuracies, the introduced dictionary approach greatly reduced the time it took to write full sentences.**

I. Introduction

Locked-in syndrome is a condition in which a person becomes unable to move or talk. While being unable to communicate through usual means, the person is still aware of the surroundings, and can typically move their eyes. To allow such a person to communicate without much help from others, a brain-computer interface (BCI) is a viable option. A BCI is a system that acquires and processes electroencephalographic (EEG) signals from the brain and transforms them into commands to control an external (electronic) device.

BCI spelling systems allow a person to write text, typically on a PC, through different types of responses detected in EEG. Responses used in spelling systems are primarily event-related potentials (ERP) related to decision-making [1][2], and visually evoked potentials (VEP) triggered through visual stimuli [3]-[5].

The new BCI system is designed at the Technical University of Denmark (DTU); hereafter we name it as the DTU BCI Speller. It is based on the steady state visual evoked potentials (SSVEP), which are detectable when a person is stimulated visually with flickering light at a fixed rate. If the response in EEG is noise-free, a simple conversion from time domain to frequency domain is enough to see which frequency the person is looking at. The advantage of SSVEP systems is that these do not require individual training of classifiers. The overall purpose is to create a system that allows a person to write arbitrary sentences without any calibration, and to have a writing speed that the person does not consider as too slow and inaccurate.

II. Methods and Materials

A. Experimental Setup

During experiment sessions, only the experimental supervisor and the test subject are sitting in an unshielded room. Inside the room, the lights are off during the experiments and the test subject is seated 60cm away from the liquid crystal display (LCD) showing stimuli.

The LCD is a BenQ XL2420T 24'' set to a refresh rate of 120 Hz. Contrast and brightness are set to maximum, resulting in a display brightness of 350 cd/m^2. The resolution is 1680x1050 pixels. Targets presented to the subjects have an area of 2.89 cm^2. The stimuli application is developed in Microsoft Silverlight and is running on a Windows 8 PC.

Three gold plated electrodes are placed along the test subject's scalp using locations from the international 10-20 system for electrode placement. The ground electrode is placed at F_{PZ}, reference electrode at F_Z and a signal electrode at O_Z. Impedances are kept around 5kΩ or lower. The amplifier used is the g.USBamp from g.tec (Guger Technologies) set to sampling rate of 512 Hz and an analog band pass filter from 5 Hz to 30 Hz. The experimental setup is illustrated in Figure 1.

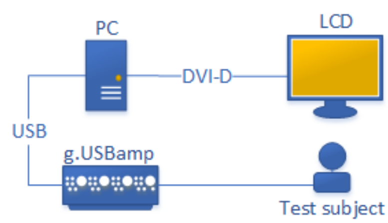

Figure 1. Experimental Setup

B. DTU BCI Speller

The user interface consists of two areas with flickering targets split by a textbox. Only one side is flickering at any given time. Below the textbox is another, always-active flickering target, the switch target, which is responsible for

[1] Technical University of Denmark, Department of Electrical Engineering, Building 349, Oersteds Plads, 2800 Kgs.Lyngby, Denmark.
[2] Department of Clinical Neurophysiology, Copenhagen University Hospital Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark.
[3] Department of Odontology, University of Copenhagen, Noerre Allé 20, 2200 Copenhagen, Denmark.

Figure 2. DTU BCI Speller interface. The left figure shows the interface when a user is selecting characters. The right figure shows the interface when a user is selecting a dictionary word.

switching between the flickering sides, see Figure 2.

The seven targets on the left side of the textbox represent a two-stage model for selecting individual characters. In the first stage, the user selects a subgroup of characters, and in the second stage, the user selects the desired character. The right side represents the dictionary with five different word targets. Each target represents a different word, and all words are updated whenever a character is written or deleted. Even though the system targets the Danish market, it supports dictionaries in both Danish and English. The different dictionaries allow direct comparison with other systems.

At any given time, there are either eight or six active flickering targets including the switch target. Since the size of each target is only 2.89 cm^2, it barely covers the fovea. The distance between any two targets is at least 1.7 cm in any direction, so that at any point, fovea can only cover one target. The used stimulation frequencies are 6 Hz, 6.5 Hz, 7 Hz, 7.5 Hz, 8.2 Hz, 9.3 Hz, 10 Hz, and 11 Hz. When a target is selected, it turns green for a brief moment, to let the user know which target is recognized. This reduces how often the user switches gaze between the textbox and individual targets. If the selected target is a word from the dictionary, a space character is added after the word, and flickering is switched back to individual characters.

C. Classifier

The classifier has two sets of data that are examined in each iteration. The duration of an iteration is approximately two seconds. The data sets are:

- *SData*: Most recent two seconds of EEG.
- *CData*: A concatenation of up to three most recent sets of *SData*.

After sampling for two seconds, autocorrelation is applied on *SData* to reduce the noise. Then FFT is applied on both sets with necessary zero-padding to obtain a frequency resolution of 0.1 Hz. Next, the classes are generated for both sets. Each class represents a target frequency. The value of each class, C_x, is the sum of power amplitudes, $|Y|$, around the relevant frequencies.

$$c_x = \sum_{H1-0.1}^{H1+0.1} |Y| + \sum_{H2-0.1}^{H2+0.1} |Y|, \qquad (1)$$

where H1 is the fundamental frequency presented, and H2 is the second harmonic. The second harmonic is taken into account, because early tests showed, that a person can have a stronger or equal response in the second harmonic as the fundamental frequency. This occurrence appears to be related to the accuracy and precision of stimulus generation.

The values in all classes are normalized in respect to each other. The dominating class will have a value of one, but the selection only happens if at least one of three quality tests is satisfied:

- The second greatest value in *SData* < 0.35.
- The second greatest value in *CData* < 0.45.
- The same class is dominating in four consecutive iterations.

The two thresholds are determined through empirical testing. Increasing the thresholds can improve selection times for some users but at the same time reduce accuracy for others.

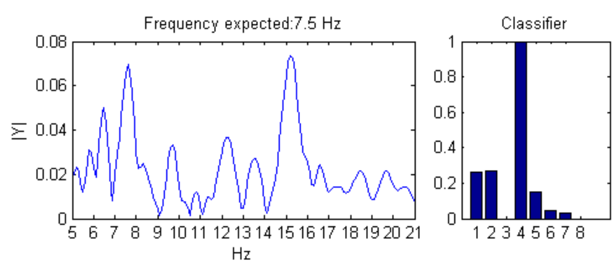

Figure 3. Classification after two seconds. Left figure represents FFT after autocorrelation. In the right figure, x-axis denotes the class and y-axis denotes the amplitude.

Figure 3 shows an example of a successful classification done after two seconds on signal where classification is not immediately evident. Looking only at the fundamental frequencies, 7.5 Hz (class 4) does not appear much larger than 6.5 Hz (class 2). However when combining the frequencies with their second harmonics, one sees that 13 Hz is not present, causing class 4, the class representing 7.5 Hz, to stand out significantly.

D. Evaluation of the BCI Spelling System

To test the system, each test subject had to write four sentences (three Danish and one English). Question marks and spaces are counted as characters. A sentence is not finished until it is correct, so any spelling mistakes along the way have to be corrected. After each sentence, the user takes a small break of less than a minute. The four sentences are:

- S1: The quick brown fox jumps over the lazy dog
- S2: Jeg vil gerne se en film
- S3: Hvad har du lavet I dag?
- S4: Zebraen ønskede sig sæbespåner

III. Results

Nine healthy subjects participated and successfully wrote all four sentences. Six males and three females, age 26.8 ± 5. Only one test subject was familiar with the concepts of BCI systems. TABLE I. shows the total amount of selections required to write all four sentences, the average time a selection takes, and the accuracies throughout all selections.

The time it takes to write a sentence is significantly lower when the subject uses the dictionary. As an example, the two subjects (4 and 7) who are fastest at selecting times had very different approaches. Subject 4 was very aware of which words were in the dictionary, while subject 7 paid little attention to it. When questioned, Subject 7 replied that the BCI responds fast enough so dictionary aid was not necessary.

TABLE I. Individual performances

Subject	***Total Selections***	***Avg. Selection time (s)***	***Accuracy (%)***
1	206	6.71	94.08
2	222	6.32	92.11
3	196	6.58	92.27
4	173	5.28	97.13
5	270	7.27	88.83
6	285	8.12	86.54
7	238	5.48	92.02
8	304	8.04	83.27
9	260	5.79	91.09
Mean ± std	239.33 ± 43.77	6.62 ± 1.03	90.81 ± 4.11

The performance of BCI Systems is usually evaluated based on the information transfer rate (ITR) expressed as bits/min. ITR is derived from the time it takes to perform a task, accuracy of the system and the amount different tasks that can be performed [6]. Figure 4 illustrates individual ITRs of the subjects for each sentence. The lowest and highest achieved ITR are 11.58 bits/min and 37.57 bits/min, respectively. It is important to note that the individual performance does not vary much, showing the robustness of the system.

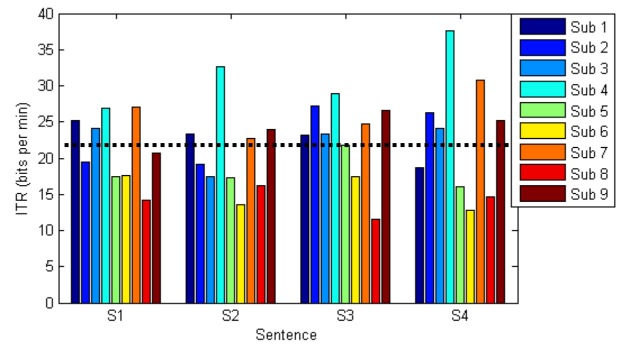

Figure 4. Individual ITR for each sentence. Dashed line shows average ITR of 21.94 bits/min.

BCI spelling systems often require more than one task to write a character. To provide a more comprehensive measure, some studies also report the amount of characters written per minute (CPM) [2][4]. Figure 5 shows all the CPM results. On average, the test subjects wrote 4.91 CPM. In the worst case, a subject wrote between 2.2 CPM and 3.4 CPM. In the best case however, another subject managed to write between 6.45 CPM and 8.74 CPM.

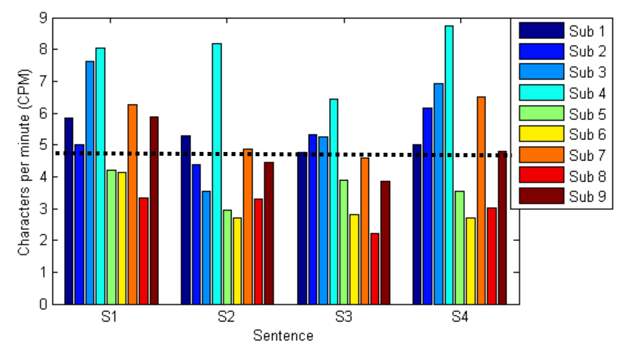

Figure 5. Individual CPM for each sentence. Dashed line shows average CPM of 4.91.

Figure 6 shows the total time it took for each test subject to write the sentences. The labels along the x-axis show the amount of characters that each sentence contains. The graph gives an intuition of how long it can take to write sentences of different lengths.

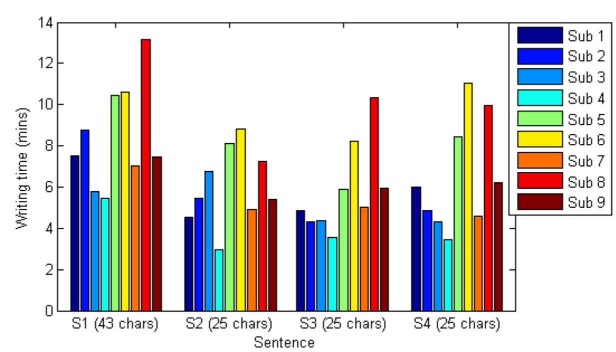

Figure 6. Individual times for writing full sentences.

IV. DISCUSSION

The system has been tested in real life conditions with healthy untrained test subjects, of which only one subject (2) was familiar with BCI systems. The system was also successfully tested in normal light conditions with lights on.

For some test subjects, the writing speed can easily be improved by increasing the threshold values. By doing so could however cause other subjects to become unable to use the system. To overcome this, a future improvement can be a more adaptive system, which based on deletions and spelling can adjusts the thresholds.

In terms of user friendliness, the system received positive feedback for both the physical setup and the usability of the software. Since only three electrodes are used, it takes approximately five minutes to place electrodes and obtain impedances below 5kΩ. The procedure is easy enough that no trained personnel are required to attach the sensors, and even relatives to a disabled user can be taught how to attach them.

Throughout the study, strong emphasis was put on writing full sentences. Aside from testing the system under real life conditions, it was done to test how exhaustive it is for a person to use the BCI over time, since SSVEP systems are known to be tiring to use and cause discomfort [8][9]. None of the test subjects felt discomfort during the session. Only subject 6 and subject 8 reported slight tiredness because their selections sometimes took more than twenty seconds to recognize. A plausible reason is that an LCD on maximum brightness only produces flickering from the targets, whereas lower brightness activates pulse-width modulation (PWM), causing additional background flicker.

The dictionary improved the amount of CPM for all but one test subject (8). On average approximately four selections were skipped each minute with the use of dictionary. Further improvement is possible by limiting the size of the dictionary. The English dictionary contains approximately 6100 words, which caused the desired words to appear fast. The Danish dictionary on the other hand consists of corpus with over 164,000 words of which subjects did not recognize many that appeared.

The flickering frequencies have a resolution of 0.1 Hz to ensure that they do not overlap in any harmonics up to 30 Hz. A downside of this resolution is that zero-padding needs to be applied. Since only the first and second harmonics are taken into account, the resolution can be lower. Previous experiments with a hardware stimulus, showed responses even in the third harmonic for some subjects. Once the software interface was implemented however, the amplitude of the third harmonic in the power spectrum was considerably lower. The cause for this occurrence is unknown as many factors have been changed, including brightness and precision of the produced flickering.

V. CONCLUSION

With this paper, a promising novel BCI spelling system, namely the DTU BCI Speller is introduced. It differentiates itself from other systems through its design and signal processing approach. The method used for SSVEP detection has quality tests that ensure high accuracy without penalizing the selection times. It also prevents selections from happening when the user looks away from the monitor, because the signal in that case is very noisy.

The developed system is user friendly, in that all test subjects felt that the interface is intuitive and it is simple to connect a person to the system. In future, the goal is to reduce the long selection times through visual feedback for the user.

REFERENCES

[1] M. S. Treder, N. M. Schmidt and B. Blankertz, “Gaze-independent brain–computer interfaces based on covert attention and feature attention,” *Journal of Engineering,* vol. 8, no. 6, p.066003, 2011.

[2] F. Aloisea, P. Arico, F. Schettini, A. Riccioa, S. Salinarib, D. Mattiaa, F. Babilonia and F. Cincottia, “A covert attention P300-based brain–computer interface: Geospell,” *Ergonomics,* vol. 55,no. 5, pp. 538-551, 2012.

[3] G. Bin, X. Gao, Y. Wang, Y. Li and B. Hong, “A high-speed BCI based on code modulation VEP,” *Journal of Neural Engineering,* vol. 8, no. 2, p. 025015, 2011.

[4] H.-J. Hwang, J.-H. Lim, Y.-J. Jung, H. Choi and S. W. Lee, “Development of an SSVEP-based BCI spelling system adopting a QWERTY-style,” *Journal of Neuroscience Methods,* vol. 208, no. 1, pp. 59-65, 2012.

[5] I. Volosyak, A. Moor and A. Gräser, “A Dictionary-Driven SSVEP Speller with a Modified Graphical User Interface,” *Lecture Notes in Computer Science*, vol. 6691, pp. 353-361, 2011.

[6] H. R. D. J. M. a. G. P. Jonathan R. Wolpaw, “EEG-Based Communication: Improved,” *IEEE Transactions on Rehabilitation Engineering,* vol. 6, no. 3, pp. 326-333, 1998.

[7] C. Neuper, G. Müiller-Putz, R. Scherer and G. Pfurtscheller, “Motor imagery and EEG-based control of spelling devices and neuroprostheses,” *Progress in Brain Research,* vol. 159, pp. 393-409, 2006.

[8] F. Beverina, G. Palmas, S. Silvoni, F. Piccione & S. Giove, “User adaptive BCIs: SSVEP and P300 based interfaces,” *PsychNology Journal,* vol. 1, pp. 331-354, 2003.

[9] N. Yoshimura and N. Itakura, “Usability of Transient VEPs in BCIs,” in *Recent Advances in Brain-computer Interface Systems*, In-Tech, 2009, pp. 119-133.