Development of brain-based smart-home/typing system

Aiana Yergaliyeva, Arnur Berikbol, Arsen Seiilkhan

April 19, 2024

1. Executive summary

Our project uses a combination of electroencephalography-electrooculography (EEG-EOG) signals to develop a highly accurate and efficient VR (Virtual reality) and MR (Mixed reality) based Brain-Computer Interface (BCI) spelling system. We aim to develop a hybrid speller system that allows users to spell words, phrases, or sentences using only their brain activity by utilizing multi-modal signals and various classification strategies. This project is primarily driven by the desire to improve the quality of life and independence of those with motor disabilities, including those suffering from spinal cord injuries, ALS, locked-in syndrome, and the elderly, by offering a substitute channel of communication. Our BCI typing system promises to greatly improve their quality of life by restoring their capacity to communicate ideas and interact with the outside world through brain activity. This will promote more independence and involvement in social, educational, and professional activities.

After reviewing the literature and relevant research, we determined that the BCI speller is one of the most well-known applications in the field. BCI spellers are based on three main paradigms: motor imagery (MI), steady-state visually evoked potential (SSVEP), and P300. Hybrid approaches combine these paradigms to achieve better performance. Because of its efficacy and efficiency, our project primarily focuses on implementing a P300 Row-Column (RC) speller. In order to overcome significant challenges and enhance the abilities of BCI spellers, we have opted to utilize convolutional neural network (CNN) classification techniques, which are well-known for their exceptional results in earlier research. Furthermore, we integrate mixed reality glasses for visual display in an effort to alleviate problems related to system size, portability, and user fatigue. This tactical change attempts to improve user comfort and maximize the quality of the EEG signal, which will improve overall system performance and character recognition.

Our project methodology includes a comprehensive experimental approach, involving environment selection and configuration, experimental setup and data collection, signal acquisition and preprocessing, data analysis for signal identification, feedback gathering on signal acquisition process, validation and iterative refinement, documentation and future development. We aim to create a catalyst for accessibility and inclusivity by embracing the revolutionary potential of virtual and mixed reality interfaces and pushing the boundaries of BCI technology. We work to create a world where communication is limitless and people of all abilities can connect, create, and prosper. We do this through innovation, teamwork, and a relentless dedication to excellence.

2. Introduction

Our project is the development of a highly accurate and fast mixed reality (MR) based BCI spelling system that employs multi-modal electroencephalography-electrooculography (EEG-EOG) signals. Users will be able to spell out words, phrases, or sentences by selecting letters on an MR screen using their brain signals. To increase the performance, we will develop a hybrid speller system that optimally combines EEG and EOG signals with diverse classification strategies so that the user and our system can act cooperatively for optimal decision-making.

The main motivation for the development of the brain-computer interface typing (speller) system is that it provides an alternative means of communication for people with motor disabilities like spinal cord injuries, amyotrophic lateral sclerosis (ALS), locked-in syndrome (when a person is only able to move their eyes) and old aged people by allowing them to convey their thoughts using their brain activity (without any physical movement). Thus, BCI typing systems increase the quality of life for such people by allowing them to engage with the world, participate in social, educational, and professional activities, and increase their independence from caregivers.

3. Background and Related Work

BCI speller is one of the most popular BCI applications. It allows people to type words via brain signals without any muscular movement. BCI speller consists of three main components: an EEG brain waves collecting device, a user interface - a grid of characters displayed on a screen, and software responsible for data processing and returning feedback [1]. The three main EEG paradigms used by BCI spellers are motor imagery (MI), steady-state visually evoked potential (SSVEP), P300, and hybrid (a combination of two or more paradigms).

In SSVEP each character flickers at a specific frequency. The frequency of the recorded EEG wave will match the frequency or harmonics of the frequency of the desired character.

The advantage of an SSVEP-based BCI speller is that no preliminary training is required for users, the disadvantages are visual fatigue caused by multiple frequencies present in stimuli and poor performance when the frequencies are close to each other or are the harmonics of each other [2]. P300 speller operates based on an ERP (event-related potential) that appears in brain waves as a positive deflection within around 300 milliseconds after the target event.

Based on the flashing pattern P300 speller divides into row-column (RC), a single character (SC), and region-based (RB). Based on the number of layers P300 speller can be uni-modal (a single layer), bi-modal (2 layers), or multi-modal. In the RC flashing pattern an entire row or a column of the matrix flashes randomly. P300 signal is elicited when the desired character's row or column is flashed. Finally, the desired character is determined from the cross-section of the row and column. In the SC flashing pattern, a single character is flashed. To complete one round of flashing in a 6x6 matrix RC speller requires 12 flashes, while the SC speller 36.

Figure 1: SC flashing pattern

Source: Adapted from [2]

Figure 2: RC flashing pattern

Source: Adapted from [2]

Contrasted to the above two patterns which are uni-modal, the RB flashing pattern requires 2 layers. In the first layer, the region containing the desired character is selected, in the second the desired character itself is selected from the selected region.

Figure 3: RB flashing pattern

Source: Adapted from [3]

For our project, we are going to implement a P300 RC speller.

BCI speller performance is measured by two criteria: accuracy - the ratio of the correct selections overall selections, and Information Transfer Rate (ITR) in bits per minute. To improve performance several innovations were proposed. One possible way of improving classification accuracy is by integrating the rapid serial visual presentation (RSVP) technique into the SSVEP speller. Character recognition can be enhanced by adding the audio stimulus to the RB speller. To evoke a stronger cortical response human facial expressions (emojis) can be used in visual stimuli. Du et al. showed that 3D visual presentation of the stimuli produces a better class discriminant of ERP features than traditional 2D presentation [3].

Different EEG paradigms can be combined in order to improve the performance resulting in a hybrid BCI speller. For instance, in simultaneous pairing, both SSVEP and P300 signals are collected and analyzed simultaneously to find the desired character. However, the disadvantages of this approach are that it might cause even more fatigue to the user and a higher computational load on the system [2].

BCI speller operates in the following steps: signal acquisition, signal preprocessing, feature extraction, classification, and output [2].

The feature extraction techniques that are widely used in BCI speller systems can be categorized as temporal (uses temporal variations of the EEG signals as a feature), frequency-based (these features provide frequency domain information of the EEG signal), time-frequency domain-based (short-time Fourier transform (STFT), wavelet transform (WT)), and deep learning-based (give detailed information about the signal, but a lot of training data is required to train the model as well as high-performance computing system) [2].

The feature optimization techniques that are widely used in BCI speller systems are principal component analysis (PCA), Fisher discriminant analysis (FDA), independent component analysis (ICA), genetic algorithms, and particle swarm optimization (PSO).

In the majority of BCI research supervised classifiers are used for classification: linear discriminant analysis (LDA), multilayer perceptron backpropagation neural network (MLP-BP-NN), and support vector machine (SVM) are the most popular ones. For our project, we are going to use the convolutional neural network (CNN) classification algorithm due to its good performance: accuracy ranging from 92% to 98% in previous studies [4].

The current open issues of BCI speller systems include the big size and low portability of the system. Also, the continuous focus on the monitor creates fatigue for the user, which affects the quality of EEG signals and diminishes character recognition. Our project will try to solve these issues to some extent by using mixed reality glasses instead of computer screens for displaying the visual paradigm.

4. Project Approach

Our project uses a phased approach to develop and validate our Brain-Computer Interface (BCI) spelling system, utilizing a variety of environments such as desktop, virtual reality (VR), and mixed reality (MR). The approach occurs as follow:

1. **Environment Selection and Implementation:** In desktop, virtual reality (VR), and mixed reality (MR) systems, we begin by choosing and implementing the BCI system. Each environment has been configured to make it easier to receive the signal and give the subjects a comfortable and immersive experience.

2. **Experimental Setup and Data Collection:** By choosing flashing symbols in each environment that correspond to predefined words, the subjects involved with the BCI system. Each flashing symbol is represented synchronously with the collection of EEG-EOG signals. This application's primary goal is to gather signals pertaining to symbol selection for later analysis.

3. **Signal Acquisition and Preprocessing:** Specialized equipment is used to record EEG-EOG signals, ensuring the accuracy of data collection. Preprocessing is done on the received signals to remove baseline bias, noise, and artifacts while preserving the signal data integrity for subsequent analysis.

4. **Data Analysis for Signal Identification:** The gathered EEG-EOG signals are examined to find patterns related to the selection of a symbol. Through the comparison of specific signal patterns with the symbols chosen by the subjects, our goal is to determine the unique neural signatures associated with each symbol.

5. **Feedback Gathering on Signal Acquisition Process:** The subjects report their experiences with the signal reception process at the end of the data collection phase. Feedback helps improve the experimental setup for subsequent iterations by focusing on factors like comfort, ease of use, and overall satisfaction.

6. **Validation and Iterative Refinement:** By using cross-validation techniques and thorough analysis, the detected signal patterns are confirmed. By continuously enhancing the signal collection procedure including modifying electrode placement, stimulus presentation, and experimental protocols any discrepancies are removed.

7. **Documentation and Future Development:** A thorough documentation of the signal identification results is provided, which offers important information for upcoming advancements. Based on the discovered neural signatures, a dependable BCI spelling system will be developed and put into use using these data.

By using this strategy, we anticipate gathering precise EEG-EOG signals associated with character selection, establishing the groundwork for the eventual creation of an accurate and efficient BCI spelling system. By taking a methodical approach, the signal collection process is guaranteed to be optimized in order to capture the neural activity required for brain-computer interfaces to function as effective means of communication.

Functional requirements:

The system should provide a user-friendly interface showing a grid of letters and a sentence that is being spelled.

The system should collect brain signals, such as EEG and EOG from a user.

The system should preprocess signals to remove noise.

The system should identify the user’s intention (letter in mind) by classifying brain signals.

Non-functional requirements:

The system should have a 90% or higher spelling accuracy.

The system should have an information transfer rate of 30 bits/min or higher with a minimized number of channels (less than 10 channels).

The system should be robust and reliable.

The system should be secure - the user's data protected from unauthorized access.

The system should be safe for use.

Domain requirements

The system should be medically certified if to be used by people with disabilities

5. Project Execution

Sixth softwares was developed for the project, 2 for each device, Desktop, Virtual reality (Meta Oculus Quest 2, further VR), Mixed Reality (Microsoft Hololens, further MR). One of them was a spelling system for “PP”, second for the word “WORLD_UNIVERSITY” word. The words were chosen to have the same number of characters. (Please see applications demonstration in additional material and/or in presentation).

These softwares are developed in order to show a user number of triggers (in each row and column) on a spelling system in order to acquire EEG and EOG signals of a human's reaction when the targeted letter is triggered/non-triggered. Total number of triggers for each of the words were 960 - 12 triggers for a character, repeated 5 times, 16 characters in total.

The EEG and EOG signals were acquired by inviting subjects, paying them for their participance and getting a consent from them under Professor Minho Lee’s Research Project. A total number of 10 healthy NU student’s were invited to the experiment that consists of two sessions.

Experiment Procedures

Subjects were invited to the lab the day before the main experiment in order to check their eyesight, teach them how to use VR and MR devices, instruct them about the procedures of the experiment, and ask them to sign the consent form.

At the beginning of the main experiment, the subjects were comfortably seated in the chair. Then, the EEG electrodes caps were worn by the subject, and 16 electrodes were placed. The Ground electrode was placed at the Fz place. Reference electrode was placed at the mastoid region of the ear. (Figure 4.)

Figure 4.

After every electrode was placed and was poured with electro-gel, the impedance was checked using ActiCap software. After that, the eeg signals were checked to have no noise using BrainVision Recorder at 100 Hz sampling rate.

The same software was used in order to record all of the 960 triggers for each of the device-word-session tuples. (total number of recording for experiment - 12, 1st session - 2 for desktop (CS, WU), 2 for MR (CS, WU), 2 for VR(CS, WU) , same for the second session)

The sessions had a 15 minute break. After each recording is collected, it is checked to have a proper number of triggers recorded(960) using the matlab code. The graphics of eeg signals on targeted character vs non-targeted character is averaged and drawn on [-20 ms, 100ms] of trigger's time interval were drawn for Oz and Cz channels to see if the data is close to truth. (Figure 5) After each pair of recordings the subject was asked to take a self-assessment questionnaire.

Figure 5. Blue - the targeted character's EEG signals, Red - non-targeted character's EEG signals.

6. Evaluation

We evaluated our system along two criteria: which paradigm the users prefer the most and which paradigm produces the best result. For that, we used self-reported questionnaires and compared signal quality across paradigms respectively.

We had two types of questionnaires: symptom questionnaires were filled after each paradigm change and consisted of 8 questions which measured the subject's physical state (level of back, neck, shoulder pain), mental state (headache), eye state (eye strain and clarity) and fatigue level (eye tiredness and drowsiness). The user experience questionnaires, which consisted of 7 questions were filled after a single full phase ended.

Figure 6: User experience questionnaire results

As Fig. 6 demonstrates the system received positive scores in terms of user experience. The users view the experience of using the system as comfortable, enjoyable, not annoying (80% of subjects), immersive, and not nauseating. They would like to experience the system again and their expectations of the system were met.

Figure 7: Symptom questionnaire results after using a desktop paradigm

Figure 8: Symptom questionnaire results after using the MR paradigm

Figure 9: Symptom questionnaire results after using the VR paradigm

As it can be observed from Fig. 7-9 there is no significant difference in scores between different paradigms.

7. Conclusion and possible future work

In conclusion, our project aims to develop a highly accurate and efficient Brain-Computer Interface (BCI) spelling system utilizing EEG-EOG signals, with a focus on enhancing the communication capabilities of individuals with motor disabilities. With the integration of VR and MR systems, we strive to create a user-friendly interface for individuals with motor impairments for spelling words, phrases, sentences using EEG-EOG signals exclusively.

While the project is ongoing and data collection is still in progress throughout April and May, the preliminary stages have shown promising results. By inviting 10 healthy NU students to participate in our experiments, we have acquired valuable EEG and EOG signals for further analysis. However, it is important to note that the larger the number of subjects involved, the more accurate the collected data will be, thus we anticipate further improvements in data quality as the project progresses. This dataset is itself a crucial step in assistive technology development for people with motor impairments.

Moving forward, the dataset will be analyzed and used for target character classification to complete the speller system. This system aims to empower individuals with disabilities and promote greater inclusivity in society.

As for further work, the dataset and the project report could serve as examples for EEG signal data collection and analysis. The EEG signals collection could lead to advancements in brain-computer interfaces, facilitating projects such as smart home control, and VR or MR applications using EEG signals exclusively.

8. References

[1] S. S. Alrumiah, "A Review on Brain-Computer Interface (BCI) Spellers: P300 Speller," Bioscience Biotechnology Research Communications, vol. 13, no. 3, pp. 1191–1199, Aug. 2020, doi: https://doi.org/10.21786/bbrc/13.3/31.

[2] S. Kundu and S. Ari, "Brain-Computer Interface Speller System for Alternative Communication: A Review," IRBM, Jul. 2021, doi: https://doi.org/10.1016/j.irbm.2021.07.001.

[3] J. Du, Y. Ke, L. Kong, T. Wang, F. He, and D. Ming, "3D Stimulus Presentation of ERP-Speller in Virtual Reality," Mar. 2019, doi: https://doi.org/10.1109/ner.2019.8717164.

[4] T. Fang et al., "Recent Advances of P300 Speller Paradigms and Algorithms," Feb. 2021, doi: https://doi.org/10.1109/bci51272.2021.9385369.

[5]: O. -Y. Kwon, M. -H. Lee, C. Guan and S. -W. Lee, "Subject-Independent Brain–Computer Interfaces Based on Deep Convolutional Neural Networks," in IEEE Transactions on Neural Networks and Learning Systems, vol. 31, no. 10, pp. 3839-3852, Oct. 2020, doi: 10.1109/TNNLS.2019.2946869.