A Thousand-Word Speller: a novel approach for communicating mental imagery through BCIs

Romero, Esteban · A01639031@tec.mx
Pulido, Miguel Ángel · A01639043@tec.mx
Sánchez, Diego · A00573206@tec.mx

All authors are affiliated to
Instituto Tecnológico y de Estudios Superiores de Monterrey · Guadalajara, México

Abstract

We developed a brain-computer-interface that aims to improve communication of mental imagery by constructing a brand new image based on a user's selection of visual input. It is based on the Unicorn Speller, a P300 speller that acquires electroencephalographic (EEG) signals and uses the P300 paradigm to select an item out of flashing rows-and-columns. Our system improves on traditional P300 speller by using an image-populated board that encompasses different categories, such as setting, subject, color, and visual style. It leverages state-of-the-art generative language and image models to combine the images' descriptions into a prompt that creates a brand new picture.

Introduction

BCIs based on EEG provide a useful form of communication in cases of motor disability where a patient can't move a part of, or even their entire body, as it occurs with lock-in syndrome. P300-based spellers have been developed along with other paradigms to ease patient communication. It uses an oddball paradigm to provide a constant stimulus and then measure and characterize the response 300 ms later to an infrequent (oddball) stimulus. Individual characters are laid out in a board-like grid pattern whose rows and columns flash with colors or images to elicit this response [1]. Currently, an advanced non-invasive P300 EEG-based speller can allow communication at up to 12 characters per minute [2].

Aphasia is an impairment in language where a patient may not be able to understand, process or produce language. It may be caused by motor impairment, degeneration, or damage to the areas that process and communicate language [3]. Patients with aphasia use other skills, such as drawing [4] that help with communication and rehabilitation. Advanced invasive BCIs have also been used very recently to fully restore language in a case of motor impairment [5]. In the case of P300-based spellers as means of rehabilitation for post-stroke motor aphasia, they have proven to be feasible means of communication by adapting them to particular needs [6].

The use of BCIs has not been limited to the field of science, but also extends to arts as tools of introspection, feedback, visual, and auditory creation [7]. We decided to create a P300 based tool to generate visual artwork by combining images. This tool may provide a faster way to communicate, sketch thoughts, ideas or generate complex artworks with the assistance of other digital programs [8]. It may be an important tool for patients with locked-in syndrome, motor disabilities and aphasia to communicate visual imagery.

Materials and Methods

The Unicorn Hybrid Black (g.tec medical engineering GmbH, Austria) is a wearable 8-channel EEG headset for BCI applications. Wet electrodes were placed over the motor cortex, and parietal and occipital lobes, namely, Fz, C3, Cz, C4, Pz, PO7, PO8, and Oz, with the reference over the mastoids. The Unicorn Speller allows the user to customize from different configurations of the P300 interface, among which we choose

the following (see fig.1): grid size of 6 rows and 10 columns; categories of image style, color, place, objects, people, and stop signal (used to stop the program); usage of the random row-column covering with famous faces paradigm; a training stage that consisted of 20 repetitions of each element, using from 4 to 5 elements to test and calibrate the classification model.

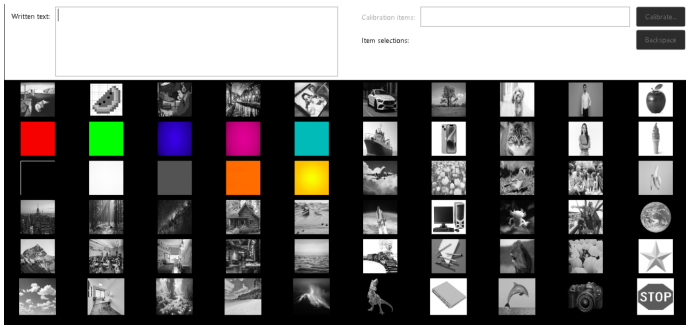

Fig. 1 Custom Unicorn Speller grid

Messages were fed into a Python program that used a dictionary to assign descriptions to each image. We used Stable Diffusion, due to its open-source nature and potential for using custom and local models, as the means for text-to-image generation through the Stability API [9] with a recent model (SDXL v2-2-2). For image prompt generation, we used GPT [10] through OpenAI's chat completions API [11]. We used the GPT-3.5 model with a contextual system message asking for the model to output an efficient DALL·E prompt based on the parameters provided by the speller's output. The final system was tested as a proof-of-concept by conducting internal testing within members of the development team to verify its functionality. No survey or participant data was collected from these trials.

Conclusion

The results obtained from our EEG-based P300 speller system with image generation demonstrated promising outcomes (see fig. 2). Participants were able to appropriately select images and within seconds of sending the stop signal, witness the creation of new images derived from mental imagery. While the outcomes of this study are promising, there are several areas for further improvement and refinement of the system. Satisfaction surveys, engagement, generation time, setup time, and other metrics will need to be collected and analyzed in further tests for more quantifiable results. Furthermore, the board could also be separated into multiple boards that are presented sequentially to reduce visual clutter, improve accuracy, and have the possibility of changing categories depending on user input. These challenges can be tackled on future research using the software developed on this proposal as a basis. We hope these types of systems can eventually reach patients to improve communication.

"Dolphin, Beach, Sunset" "Space, Lion, Digital Art"

Fig 2. Examples of the generated images

References

[1] Pan, J., Chen, X., Ban, N., He, J., Chen, J., & Huang, H. (2022). Advances in P300 brain–computer interface spellers: toward paradigm design and performance evaluation. Frontiers in Human Neuroscience, 16. https://doi.org/10.3389/fnhum.2022.1077717

[2] Speier, W., Arnold, C., Chandravadia, N., Roberts, D., Pendekanti, S., & Pouratian, N. (2018). Improving P300 spelling rate using language models and predictive spelling. Brain Computer Interfaces (Abingdon, England), 5(1), 13–22. https://doi.org/10.1080/2326263x.2017.1410418

[3] Gutenstein, M. (1999). Talking about aphasia. BMJ (Clinical Research Ed.), 319(Suppl S6), 9912475b. https://doi.org/10.1136/sbmj.9912475b

[4] Sacchett, C. (2002). Drawing in aphasia: moving towards the interactive. International Journal of Human-Computer Studies, 57(4), 263–277. https://doi.org/10.1006/ijhc.2002.1018

[5] Metzger, S. L., Littlejohn, K. T., Silva, A. B., Moses, D. A., Seaton, M. P., Wang, R., Dougherty, M. E., Liu, J. R., Wu, P., Berger, M. A., Zhuravleva, I., Tu-Chan, A., Ganguly, K., Anumanchipalli, G. K., & Chang, E. F. (2023). A high-performance neuroprosthesis for speech decoding and avatar control. Nature, 620(7976), 1037–1046. https://doi.org/10.1038/s41586-023-06443-4

[6] Kleih, S. C., Gottschalt, L., Teichlein, E., & Weilbach, F. X. (2016). Toward a P300 based brain-computer interface for aphasia rehabilitation after stroke: Presentation of theoretical considerations and a pilot feasibility study. Frontiers in Human Neuroscience, 10. https://doi.org/10.3389/fnhum.2016.00547

[7] Prpa, M., & Pasquier, P. (2019). Brain-computer interfaces in contemporary art: A state of the art and taxonomy. In Brain Art (pp. 65–115). Springer International Publishing.

[8] Creed, C. (2018). Assistive technology for disabled visual artists: exploring the impact of digital technologies on artistic practice. Disability & Society, 33(7), 1103–1119. https://doi.org/10.1080/09687599.2018.1469400

[9] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2021). High-resolution image synthesis with latent diffusion models. https://doi.org/10.48550/ARXIV.2112.10752

[10] OpenAI. (2023). GPT-4 Technical Report. https://doi.org/10.48550/ARXIV.2303.08774

[11] OpenAI. (2022). Introducing ChatGPT [Blog]. Retrieved August 31, 2023, from https://openai.com/blog/chatgpt