SSVEP:

Using Visual Evoked Potentials as a BCI

Claire Leahy and Lute Lillo
BME 6710 / CS 6990 / NSCI 6990: Brain-Computer Interfaces Project 2 Report
University of Vermont
Burlington, VT

INTRODUCTION

Steady-state visual evoked potentials (SSVEP) are brain signals that are produced in response to a visual external stimulus [1]. As a consequence of visually observing a stimulus with a set frequency, the electroencephalography (EEG) signal will change such that the stimulus frequency may be observed in select parts of the brain. Because these stimuli produce such marked changes to the frequency spectrum of the EEG data, this spectrum may be used as a brain-computer interface (BCI) when a user is presented with multiple stimulus frequencies. In the provided code, the SSVEPs are thoroughly investigated as they are produced from unfiltered EEG data. From the frequency spectrum of the data, the code produces predictions across possible epoch times for which stimulus was present based on the magnitude of the power for either stimulus frequency. Using these predictions, the code produces figures of merit such as accuracy and ITR (information transfer rate in bits per second) to describe how well the predictions classify the stimuli. These predictions, as well as the figures of merit, are subsequently provided to offer a visual interpretation of the efficacy of classification. By investigating these plots, a reasonable epoch length (one with both a high ITR and high accuracy), channel (an electrode that depicts a high power for the stimulus frequencies), and threshold (the difference between the power at each stimulus frequency) can inform the user of the expectations of the SSVEP as a working BCI.

METHODS

To generate the predictions as well as the truth data for each epoch, the code includes a function called **generate_predictions()**. This code provides predictions both quantitatively and qualitatively, returning arrays containing the difference of the power (predictor) at both stimulus frequencies and a boolean value describing whether the higher or lower frequency stimulus possessed the higher frequency (in this case True when the predictor was positive, False otherwise). This function adapts to the frequencies present in Python's MNE SSVEP dataset, where the event_types field is extracted to determine the frequencies (in this case 12Hz and 15Hz); throughout the code, the higher stimulus represents True for a positive predictor. The classifier enables the user to generate predictions for different channels or relative start and end times within the epochs, as each of those variables is an optional parameter.

Using the labels (true and predicted), **calculate_figures_of_merit()** is a function that produces the accuracy (proportion of correctly classified results) and the ITR (a quantity that is based on accuracy and the number of stimuli). This function also stores a list of the quantitative predictor values organized by epoch when the signal was present or absent. This function is capable of performing its operations for a variable number of stimuli, as it takes an optional variable classes_count (in this case, the default is 2 to signify the binary nature of these data).

While the previously described functions only work for a single start and end time pair, **figures_of_merit_over_epochs()** loops through a variety of possible start and end time pairs over each epoch for a given dataset and electrode. This function stores the figures of merit (accuracy, ITR) from each epoch time pair as an array; for pairs within the start and end ranges given, invalid times are given placeholder values, where accuracy is 0.5 (equivalent to a classifier that randomly predicts for two stimuli) and ITR is 0.00 (ITR is 0.00 for 1/N accuracy, where N is the number of classes [2]). Furthermore, this function stores the quantitative predictors for when the signal is present as well as when it is absent in a list for the current epoch before those lists are stored in a tuple (organized by present and absent signal). This function is capable of performing these iterations over any combination of valid start times and end times (each in a unique array) for a given channel, although the default arrays contain 0-19s, as the event durations are 20 seconds.

The final two functions for this project include **plot_figures_of_merit()**, which is a function that plots the accuracy and ITR values in a pseudocolor array for each combination of start and end times, and **plot_predictor_histrogram()**, which takes a specific epoch start and end time and plots the density of the truth labels (over the range) against the threshold. Both figures have the option to indicate the subject and channel of interest, although start and end times (either array or integer) are required parameters.

Since SSVEPs are best observed in the occipital and parietal lobes [3], meaning they have notable effects on the frequency spectrum such that the stimuli frequencies can be clearly observed, it follows that the occipital electrodes (namely channel O2) were examined most closely. Each channel was briefly observed for both subjects, though the knowledge of the SSVEPs' impacts on the different lobes heavily influenced the decision to look more in depth at the occipital electrodes, as did the clear presentation of the data for these channels. After plotting the accuracy and ITR plots, an epoch time range of 5 seconds to 11 seconds was considered due to its high (but not perfect) accuracy and high ITR. Using this start and end time, the predictor quantities for present and absent signals were presented, thus permitting the selection of a reasonable threshold over a shorter period of time than the entire event duration.

RESULTS

In Figure 1, a pattern is observed with respect to the start and end times in the channel O2 for subject 1. As the difference between the end and start time increases, so does the accuracy (measured by percentage of corrected true positives), and the ITR times. This pattern becomes clearer for subject 2 as represented by Figure 2. For later start times, shorter epoch lengths are required to obtain higher accuracy, which can be represented by the narrowing of the triangle along the diagonal (Figure 1, 2).

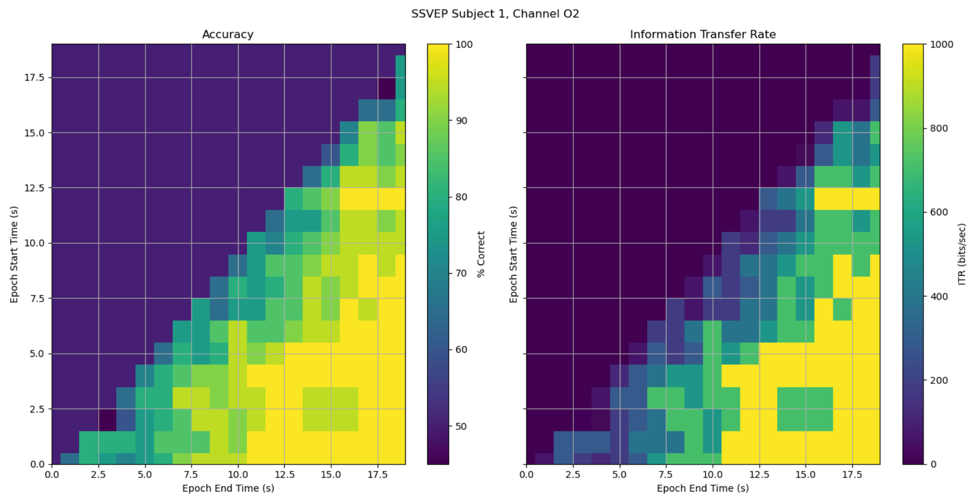

Figure 1: Figures of merit plots for subject 1 at channel O2.

Based on Figure 1 and 2, there is no need to perform the event for the total duration of these trials (20s based on the data) in order to obtain accurate results. It is possible to gather these results over a shorter period of time; for example, setting the start time at 0s and the end time at 10s depicts a high accuracy for channel O2 for both subjects (Figures 1, 2). Therefore, the epochs are comprised of 10 seconds rather than the 20 second used in the data. Examples of this trend are observed in subject 2 and in channels O1 around the 12.5s mark and channel OZ around 7.5s. However, the trend is less noticeable for the non-occipital channels, and a larger sample data of multiple subjects will be necessary to claim such a hypothesis.

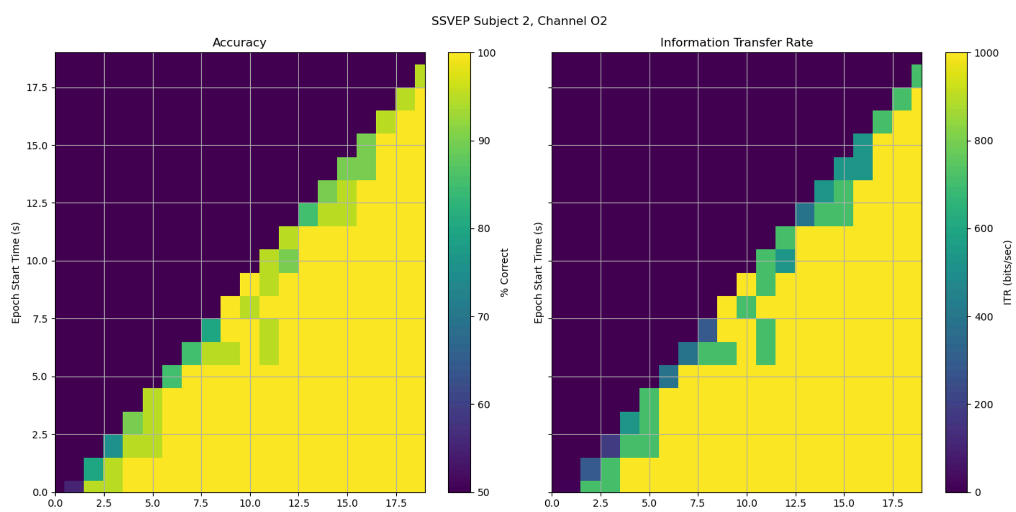

Figure 2: Figures of merit plots for subject 2 at channel O2.

Conversely, Figure 3 shows the density distribution of the predictors for cases in which the signal was either present or absent. The accuracy and ITR results have influence, or better said, can be traced back by observing at the predictor plots. For subject 2, the predictors exhibit less overlap than for subject 1; consequently, the classifier is more accurate for subject 2 over the given time range (5-11s) than it is for subject 1.

Some things to take into account are that the predictors will inherently produce some false results. While this is not necessarily something observed in Figure 3, this behavior can be assumed due to the use of a particular threshold to produce the prediction densities. Therefore, there will be an implicit bias given by the predictor of choice and/or the threshold used. Figure 3 does, however, suggest that shifting the threshold may be necessary to obtain a useful BCI, whether emphasis be placed on accuracy, specificity, or sensitivity. Presently, the threshold would place value on classifying the signal somewhat accurately and slightly specific.

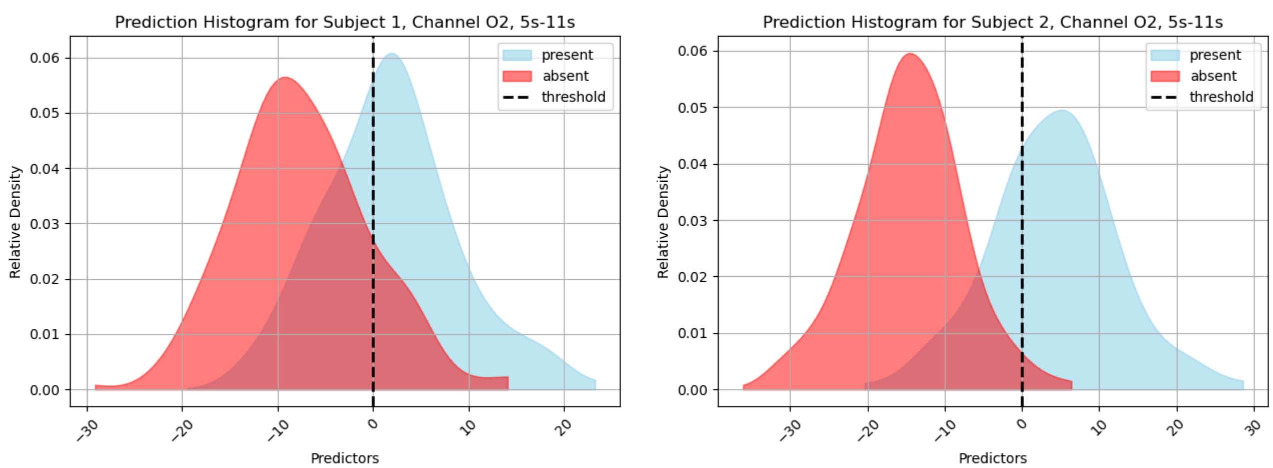

Figure 3: Prediction histograms for channel O2 over the epoch range 5s to 11s (Left: Subject 1, Right: Subject 2)..

DISCUSSION

While the greatest accuracy did frequently occur over the entirety of the event (0s start and 20s end), the entire range was often not needed to accurately classify the signal. Using an epoch time over the span of roughly 5 to 11s demonstrates high accuracy and a high ITR for the occipital channels, which are most highly implicated in SSVEP data. Because of the relatively high accuracy over the short period of time, the delayed start to the epoch will permit the user to become accustomed to the frequency while requiring less time overall (less data will need to be processed with a delayed start while the end time is well before the end of the present event durations, suggesting the following 10 seconds are not required for generating a functional BCI).

When accuracy is valued, the highest proportion of correct classifications is the measure of success. For the selected time range, this threshold would be around -4dB, as the predictor histograms depict the intersections at -3.3dB (subject 1) and -5.6dB (subject 2) to ensure the highest possible amount of true positives and true negatives, in turn minimizing false positives and false negatives. For two normally distributed datasets, as has been applied to the predictor data, this threshold falls where the two curves intersect. Conversely, sensitivity would be more highly valued in which it is important to capture all true positive cases, suggesting the need to shift the threshold leftward. For an epoch of the chosen range, the threshold would have to be around -20dB to capture every positive case. Finally, to account for specificity, the threshold would have to be shifted rightward to limit the amount of false positives. The variety is greater between the predictors for the true negatives than for the intersection or true positive points across the subjects; however, on average, the specificity threshold is about 10dB. It is important to note that the predictions themselves inherently change depending on the predictor given, so the densities would change about the various predictors.

One case in which accuracy would be the desired figure of merit would be using the SSVEP to predict "Yes" or "No" for someone unable to talk. For example, "Yes" could flash at the lower frequency, while "No" could flash at the higher frequency. Valuing accuracy enables the greatest proportion of intended responses rather than overestimating the number of either "Yes" or "No" responses, which would be desirable in the context of conversation. A motor action that would value accuracy is movement of the left extremity or the right extremity for similar reasons.

Sensitivity is valued when the consequence of a false negative is much greater than that of a false positive [4]. For example, someone who is unable to communicate without a BCI may require a high-sensitivity BCI to relay discomfort or pain to alert a nearby individual of a problem the individual may otherwise be unable to convey. Detection of a problem is incredibly important to receive proper care, so even if there is a high rate of false positives, a subsequent BCI with high accuracy (such as the "Yes" or "No" communicator) could potentially be utilized to detect whether the alert was a false alarm. While it may be an inconvenience to have many false positives, the consequence of not detecting a potential issue could be dire.

On the other hand, specificity will be important if false positives have detrimental consequences, such as causing the user harm [4]. While movement of a mobility device such as a motorized wheelchair could value accuracy, there is also an argument for specificity depending on where the device will be used. For example, within a house, accuracy may be more important to navigate generally to a desired location; however, if the device is to be used next to a road, improperly classifying "straight" as "left" could push the device into traffic, making specificity more important in the context. It is important to note that in most contexts, valuing a particular figure of merit does not need to be absolute, and an intermediate weighting may be considered.

The above results suggest that SSVEPs may be used as functional BCIs with high accuracy and ITRs when the proper channels are chosen. For SSVEPs, these electrodes are required to be located on the occipital lobe. Moreover, these BCIs can perform more efficiently than suggested by investigating accuracy and ITR of different points in the epoch, effectively shortening the length, and in turn the amount, of data required to classify the signals. Finally, the threshold intrinsically affects the predictions generated; however, it can be adjusted to match the user's classification needs.

REFERENCES

[1] D. Jangraw, “13_SSVEPs,” University of Vermont, Feb. 27, 2024.

[2] D. Jangraw, “18_SignalDetectionTheory,” University of Vermont, Mar. 28, 2024.

[3] Z. İşcan and V. V. Nikulin, “Steady state visual evoked potential (SSVEP) based brain-computer interface (BCI) performance under different perturbations,” *PLoS One*, vol. 13, no. 1, p. e0191673, Jan. 2018, doi: 10.1371/journal.pone.0191673.

[4] D. Jangraw, “19_EvaluatingPerformance,” University of Vermont, Apr. 04, 2024.