ON VISUALLY EVOKED POTENTIALS IN EEG INDUCED BY MULTIPLE PSEUDORANDOM BINARY SEQUENCES FOR BRAIN COMPUTER INTERFACE DESIGN

H. Nezamfar[1], *U. Orhan*[1], *D. Erdogmus*[1], *K.E. Hild*[2], *S. Purwar*[1], *B. Oken*[2], *M.Fried-Oken*[2]

[1] Cognitive Systems Laboratory, Northeastern University, Boston, MA, USA
[2] Oregon Health and Science University, Portland, OR, USA

ABSTRACT

Visually evoked potentials have attracted great attention in the last two decades for the purpose of brain computer interface design. Visually evoked P300 response is a major signal of interest that has been widely studied. Steady state visual evoked potentials that occur in response to periodically flickering visual stimuli have been primarily investigated as an alternative. There also exists some work on the use of an m-sequence and its shifted versions to induce responses that are primarily in the visual cortex but are not periodic. In this paper, we study the use of multiple m-sequences for intent discrimination in the brain interface, as opposed to a single m-sequence whose shifted versions are to be discriminated from each other. Specifically, we used four different m-sequences of length 31. Our main goal is to study if the bit presentation rate of the m-sequences have an impact on classification accuracy and speed. In this initial study, where we compared two basic classifier schemes using EEG data acquired with 15Hz and 30Hz bit presentation rates, our results are mixed; while on one subject, we got promising results indicating bit presentation rate could be increased without decrease in classification accuracy; thus leading to a faster decision-rate in the brain interface, on our second subject, this conclusion is not supported. Further detailed experimental studies as well as signal processing methodology design, especially for information fusion across EEG channels, will be conducted to investigate this question further.

Index Terms— Brain computer interface, electroencephalography, EEG, visually evoked potential, SSVEP, pseudorandom M-sequence

1. INTRODUCTION

Brain computer interfaces (BCI) are receiving increasing attention as a novel human computer interaction framework that could allow the communication of a person's intent to a computer-enabled application in a seamless manner. In electroencephalography (EEG) based BCI design the visually evoked P300 potential that is induced as a response to flashing stimuli (e.g., letters or icons) have been studied extensively leading to the popular P300-Speller paradigm and its variations [1, 2]. On a similar note, we have been studying the use of rapid serial visual presentation (RSVP) of pictures (of objects or letters) to induce the visual P300 by forcing a mental target matching process [3, 4]. Berlin BCI also has been converging to a paradigm that resembles RSVP, but using a two-level hierarchical hexagonal decision tree approach as opposed to our right-sided binary decision tree approach [5].

This work is supported by NSF under grants ECCS0929576, ECCS0934506, IIS0934509, IIS0914808, and BCS1027724 and NIH grant 1R01DC009834-01. The opinions presented here are solely those of the authors and do not necessarily reflect the opinions of the funding agencies.

The steady state visually evoked potentials (SSVEP) that are induced by flickering visual stimuli that follow a periodic flickering pattern, such as a checkerboard pattern that flips colors between black and white in each box at a preset frequency, have been the major contender in the VEP domain [2, 6]. In this approach, among multiple flickering stimuli, the subject focuses gaze and attention to the one that represents the desired command or action. The periodic flickering induces oscillatory activity in the visual cortex at matching frequencies (harmonics) that can be measured and detected in EEG using spectrum analysis [7, 8]. The drawback of frequency-based discrimination is the requirement for longer waveforms for improved frequency resolution; if two very close frequencies exist, the discrimination performance degrades considerably for short-sequence spectrum analyzers [9], as expected.

An alternative to periodic black-white flickering checkerboards have been the use of periodic m-sequences that have been first studied and proposed as a BCI design mechanism by Sutter two decades ago [10, 11]. In this variation of the paradigm, the checkerboard or visual stimuli are temporally flickered according to phase-shifted versions of a selected m-sequence; the approximate self-orthogonality of the shifted versions of a given m-sequence allows one to design a relatively simple BCI signal processing solution using template matching to detect one of many phases of the m-sequence. This work has been revived by Gao and colleagues in the recent years [12]. Using different lagged versions of one M-sequence allows us to have as many possible command options as the length of the M-sequence. However, if the number of desired selectors increases, it requires M-sequence to be much longer, which is not desirable.

In this paper, we study the case where multiple m-sequences [13], pseudorandom binary sequences with reasonable cross-correlation properties, are used. The advantages of using multiple m-sequences over shifted version of one sequence include the potential for signal processing and detector algorithm designs that do not require the knowledge of precise timing of stimuli - as blind multiuser detection in wireless communication works, for instance. In contrast, if one uses the same sequence and tries to discriminate between time-shifted versions, precise timing information is necessary for operation. In the current study, we also assume that timing information is available so that a template matching classifier can be employed as Bin and colleagues do [12]. We also study simple and suboptimal decision fusion techniques as a first attempt. Our main goal is to investigate if we can use different bit presentation rates without an impact on VEP classification accuracy. If we can classify VEP corresponding to different m-sequences with the same accuracy when a given m-sequence is presented at a faster bit rate, then we have the opportunity to increase BCI bandwidth by simply increasing the bit presentation rate leading to faster decisions by the BCI signal processing algorithm and faster intent detection for

BCI-controlled applications.

BCI based visual attention detection has one advantage over eye tracker based gaze detectors; by being able to detect the covert attention, which is the mental focus change without shifting the gaze one could design human computer interactions systems for persons who do not have precise gaze control [6]. People with very severe levels of motor impairment for whom intentional gaze control is a challenge, this modality becomes a promising alternative. Accuracy and speed, as usual, are the main concerns for applications relevant to this group.

2. DATA ACQUISITION

As the visual stimulus, we use two inverted checkerboard patterns with a 10×10 black-white blocks centered on the screen covering a $21cm \times 21cm$ area. The subject is seated such that the checkerboard is approximately centered in the field-of-view and the eye to screen distance is approximately 60cm away, leading to an approximate visual angle of 20°. The subjects are not restricted to maintain the visual or viewing angle during data acquisition. The binary sequence that is presented on the screen was also measured and recorded using an optical sensor synchronously with the EEG using a g.USBamp and g.TRIGbox from G.tec (Graz, Austria). The two inverted versions of the checkerboard are arbitrarily assigned the bit labels 0 and 1 and the appropriate checkerboard was sent to the screen using Psychophysics Toolbox in the first possible monitor refresh cycle. As monitor refresh rate is set to 60Hz, our frequency selections for bit presentation rate are guided by this limitation and we try 15Hz and 30Hz bit rates in order to ensure that visual stimulus transitions occur precisely at the intended times.

For this study, the m-sequence set consists of 4 elements, each one with 31-bits. The sequences are selected from among all 31-length m-sequences in order to approximately minimize the pairwise crosscorrelations. During an experimental session, for each trial one of the four sequences is selected randomly in an independent identically distributed fashion according to a uniform probability distribution. The session consisted of 80 trials and each trial contained 12 periods of the designated m-sequence. For a given session, the bit presentation rates were fixed at either 15Hz or 30Hz. Each trial begins with a one second fixation period during which the subject is instructed to focus the gaze on the $+$ sign at the center of the screen in preparation for the upcoming trial. Between consecutive trials (each of which approximately lasts 25s or 13s) the subject can rest as much as needed and initiates the next trial with a button press at will.

EEG signals, along with the optical sensor data, are captured from the scalp using active g.Butterfly electrodes using a g.Gammabox and a g.USBamp by G.tec. A nonabrasive conductive gel is used to provide conductivity between the scalp and the electrodes. Since the goal is to detect modulated P100 signals from the visual cortex, EEG sites were selected to have a higher spatial density around the visual cortex: O2, Oz, O1, PO4, POz, PO3, P4, P2, Pz, P1, P3, Cp2, Cp1, C4, Cz, C3. Table 1 shows the mapping of channel indices to their position on the scalp.

Table 1. Correspondance map between EEG electrode positions and channel indices.

Channel Index	Electrode Position	Channel Number	Electrode Position
1	C3	9	P2
2	Cz	10	P4
3	C4	11	PO3
4	CP1	12	POz
5	CP2	13	PO4
6	P3	14	O1
7	P1	15	Oz
8	Pz	16	O2

3. CLASSIFICATION METHODS

Two variations of template matching based classification is utilized in our EEG data analysis: (i) determine and use only the best EEG channel to make a decision between the candidate m-sequences; (ii) use data from all 16 EEG channels in a concatenated vector and make a single template matching decision based on this large temporal feature vector. In general, one needs to identify the correlation structure between the decisions of classifiers based on each channel and then develop a Bayesian fusion rule that is consistent with this model - a naive Bayesian approach assuming independence would be the first thing to be tried, but other models that use graphical models allowing spatially neighboring channels to show correlation are also possible. These Bayesian fusion approaches are left as future work for the moment. If done properly, fusion of information across multiple EEG channels will lead to improved classification accuracy since spatial diversity of the sensors will contribute novel information. Since our main goal is to make a decision as fast as possible, in the current study, the template matching classifiers use only and exactly one-period-long EEG traces as templates and decisions can be made in as little as one period of the m-sequences; approximately 1s for 30Hz bit presentation rate, and approximately 2s for 15Hz rate, in this study. Clearly, if the classifiers utilize longer templates, accuracy will monotonically increase at the cost of introducing longer decision delays when the user intends to switch between m-sequences (i.e. commands to the computer).

3.1. Single Channel Template-Matching

This is a correlation-coefficient-based template matching classifier, similar to the matched filter in spirit. The template has to be learned from calibration/training data collected from the visual cortex areas with EEG in response to each m-sequence. Currently, the template is taken as the sample mean of EEG responses at each given channel in response to an m-sequence period, aligning the first sample of the template to the onset sample of the m-sequence on the screen as measured by the optical sensors. Clearly, larger numbers of calibration sequence periods will make the template smoother and less noisy. Under the assumption of Gaussian background and measurement noise, the sample averaging procedure gives us the maximum likelihood template, however, in future work, to improve system robustness, we must and will investigate more robust statistical model learning techniques that will not only learn the template but also the natural variations so that outlier signals can be detected in test mode probabilistically and rejected; such outliers may occur due to muscle or other artifacts. These templates are made for each EEG channel separately. The classifier simply correlates the 4 templates built for each sequence with the EEG signal at each channel time-locked to the m-sequence transitions obtained from the optical sensor measurement. The m-sequence template that yields the highest correlation coefficient is selected. Specifically, the decision of channel c is $D_c = \arg\max_i \rho_i^c$, where ρ_i^c is the correlation coefficient between the t_i^c template for the i^{th} m-sequence for channel c and s^c the windowed EEG signal from that channel given by

$$\rho_i^c = s^{cT} t_i^c \tag{1}$$

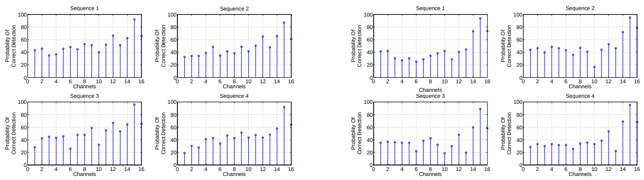

Fig. 1. Probability of correct decision for individual channels for each m-sequence using 15Hz bit presentation rate – on the left 2×2 block male subject, on the right female subject.

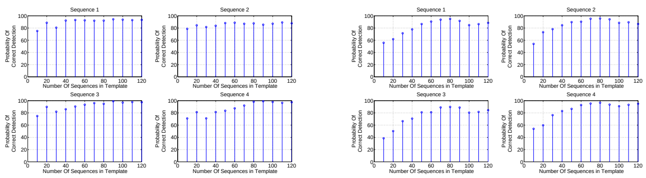

Fig. 2. Probability of correct decision on test set using different number of training samples to determine templates for Oz-channel classifier for each m-sequence using 15Hz bit presentation rate – on the left 2×2 block male subject, on the right female subject.

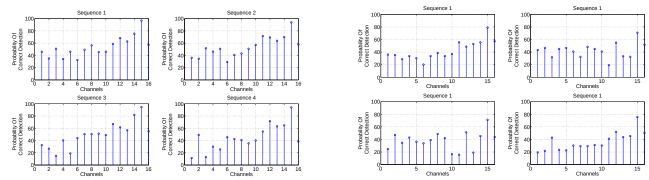

Fig. 3. Probability of correct decision for individual channels for each m-sequence using 30Hz bit presentation rate – on the left 2×2 block male subject, on the right female subject.

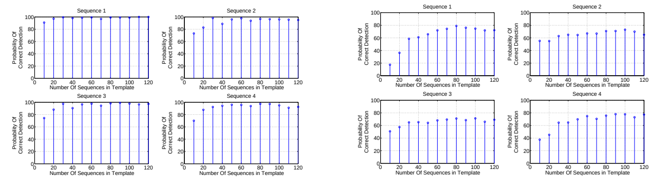

Fig. 4. Probability of correct decision on test set using different number of training samples to determine templates for Oz-channel classifier for each m-sequence using 30Hz bit presentation rate – on the left 2×2 block male subject, on the right female subject.

The performance of each channel is estimated on training data with k-fold cross-validation and the best performing channel is selected as the only one that makes the final decisions.

3.2. Multi-channel Template Matching

This classifier performs the same operation as above but concatenates the signals from all channels into a large template and signal vector. Specifically, the decision is given by $D = \arg\max_i \rho_i$, where $\rho_i = \left(s^T t_i\right)$. Here, $s = [s_1, s_2, \ldots, s_{16}]^T$ and $t_i = \left[t_i^1, t_i^2, \ldots, t_i^{16}\right]^T$.

4. RESULTS

We performed experiments using two volunteer subjects, a 22 year old male and a 27 year old female. Both subjects were healthy with normal vision. Each subject participated in 2 separate sessions. EEG acquisition was performed as described above using flickering checkerboard patterns according to m-sequences that were designed in advance. The sessions were for 15Hz and 30Hz bit presentation rates, respectively. The data from each session is split into training and test portions; out of the 200 trials in each session ($50 * 12 = 600$ periods for each of the four m-sequences), the training data was always selected as the first N_{train} periods of each m-sequence appearing in the trials. Given N_{train} and designated training data as described, remaining periods of all m-sequences were used for testing purposes to generate performance results presented in the figures.

First, we present the results for sessions using 15Hz bit presentation rate. Using $N_{train} = 70$ for each m-sequence, the template is formed for individual channels and test performance is evaluated. Figure 1 shows the probability of correct decision for different channels for our subjects. As expected, channels Oz, O1, and O2 stand out in performance, while Oz emerges as the best individual channel, being positioned right on top of the occipital lobe where the visual cortex is located. For this best channel, Oz, we investigate the effect of number of training samples on performance by generating the templates using an increasing number of training sample periods (from 10 to 120). Results shown in Figure 2 demonstrates that the performance of the template-matching classifier stabilizes after around 60 or 70 samples in both subjects for all m-sequences. Using a sufficiently large number of samples and Oz-signals alone, one can get well over 90% accuracy. Figures 3&4 present the same results for the 30Hz bit presentation rate experiments. These results indicate that for the male subject, changing the bir presentation rate did not have a significant negative or positive impact on performance -in fact, the accuracy of channel Oz increased slightly. On the other hand, the 30Hz results for the female subject using Oz only is considerably worse than her 15Hz experiment results. Based on the investigation of single channel accuracy for all channels, we suspect that the O-channels (O1, Oz, O2) have probably been improperly placed/connected and signal quality at these channels was low for this female-30Hz session. Figures 5&6 show the probability of correct decision of the second classifier for both subjects at 15 and 30 Hz.

5. DISCUSSION AND FUTURE WORK

Our limited experimental results, assuming that the poor performance of one subject at the session with higher bit presentation rate is due to poor electrode placement on the occipital lobe, could be indicative of the possibility of utilizing higher bit rates in flickering stimuli without significant influence on BCI performance. This hypothesis, must still be verified in future work using a larger set of subjects and sessions. It has been clear that not all scalp locations provide equally useful EEG signals for SSVEP classification; as expected, sites near the visual cortex resulted in better accuracy when used individually.

The m-sequences used in the current experiments were relatively short – only 31 bits long; consequently their cross-correlations could not be made extremely small. Future work will focus on determining and designing binary pseudorandom sequences with favorable cross-correlation properties that will enable us to design more accurate classifiers. This correlation improvement is likely to come at the

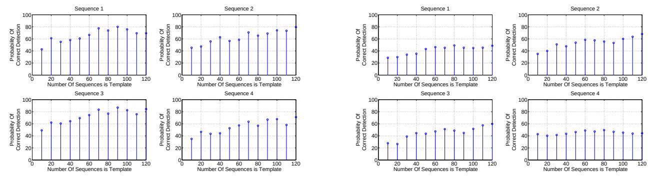

Fig. 5. Probability of correct decision on test set using different number of training samples to determine templates for the second classifier for each m-sequence using 15Hz bit presentation rate – on the left 2×2 block male subject, on the right female subject.

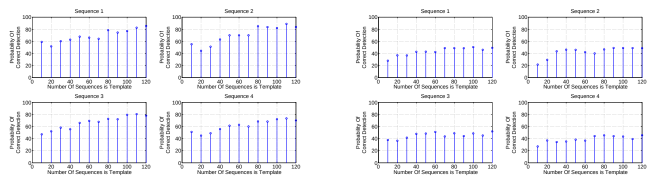

Fig. 6. Probability of correct decision on test set using different number of training samples to determine templates for the second classifier for each m-sequence using 30Hz bit presentation rate – on the left 2×2 block male subject, on the right female subject.

cost of increased sequence length; however, if our hypothesis regarding the increasing of bit presentation rate without any consequence in detector performance holds in future experiments, the additional time costs in decision making can be eliminated.

The results using even relatively poorly designed m-sequences demonstrate that it is possible to achieve over 95% classification accuracy among four classes. Better designed sequences and better designed classifiers can only improve performance. Current study focused on a very basic matched filter classifier that used one or all channels for initial assessment. In future work, we will also consider developing generative statistical signal models that will give rise to principled and statistically optimal classifiers that are based on more sophisticated models. For instance, an immediate future analysis that will be performed is the use of Bayesian fusion for channel decisions at the binary output level. Simply stacking up features from each channel into a large feature vector performed worse than using only the features from the best channel – this outcome is to be expected from a general feature dimension reduction perspective. However, besides Oz, there are other channels that carry potentially useful novel statistical information about the class label; consequently, their detection outcomes could be exploited to increase performance. In future work, we will consider Bayesian fusion models that assume conditional independence of channel-specific classifier decisions, as well as conditional dependence of these only up to a certain spatial distance (best radius to be determined from data statistically). These fusion approaches are expected to improve classification accuracy to some extent.

Training (calibration) time is a major concern in BCI system design. If BCI systems are to become a practical and widely used tool, then the calibration of the signal classifiers must not take an extremely long amount of time prior to each use. In our best case we were able to use 30 periods of each sequence to build the templates and achieve more than 90 percent accuracy in detecting the four m-sequences. For this case, the total training time will be approximately 2.5 minutes. With more sophisticated signal models that capture subject-to-subject and session-to-session variability in signal statistics, it is conceivable that the training time could be reduced significantly by exploiting information about optimal detectors found and used in previous sessions of the same subject or even different subjects.

6. REFERENCES

[1] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, and T.M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[2] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlogl, B. Obermaier, and M. Pregenzer, "Current trends in Graz brain-computer interface (BCI) research," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 216–219, 2000.

[3] S. Mathan, D. Erdogmus, Y. Huang, M. Pavel, P. Ververs, J. Carciofini, M. Dorneich, and S. Whitlow, "Rapid image analysis using neural signals," in *CHI'08 extended abstracts on Human factors in computing systems*. ACM, 2008, pp. 3309–3314.

[4] S. Mathan, P. Ververs, M. Dorneich, S. Whitlow, J. Carciofini, D. Erdogmus, M. Pavel, C. Huang, T. Lan, and A. Adami, "Neurotechnology for Image Analysis: Searching for Needles in Haystacks Efficiently," *Augmented Cognition: Past, Present, and Future,*, 2006.

[5] M.S. Treder and B. Blankertz, "(C) overt attention and visual speller design in an ERP-based brain-computer interface," *Behavioral and Brain Functions*, vol. 6, no. 1, pp. 28, 2010.

[6] B.Z. Allison, D.J. McFarland, G. Schalk, S.D. Zheng, M.M. Jackson, and J.R. Wolpaw, "Towards an independent brain-computer interface using steady state visual evoked potentials," *Clinical Neurophysiology*, vol. 119, no. 2, pp. 399–408, 2008.

[7] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain-computer interface with high transfer rates," *Biomedical Engineering, IEEE Transactions on*, vol. 49, no. 10, pp. 1181–1186, 2002.

[8] J. Mast and J.D. Victor, "Fluctuations of steady-state VEPs: interaction of driven evoked potentials and the EEG," *Electroencephalography and clinical neurophysiology*, vol. 78, no. 5, pp. 389–401, 1991.

[9] K.K. Shyu, P.L. Lee, Y.J. Liu, and J.J. Sie, "Dual-Frequency Steady-State Visual Evoked Potential for Brain Computer Interface," *Neuroscience Letters*, 2010.

[10] E. Sutter, "The Visual Evoked Response As A Communication Channel," *Proceedings of the Symposium on Biosensors*, pp. 95–100, 1984.

[11] E. Sutter, "The brain response interface: Communication through visually-induced electrical brain responses," *Journal of Microcomputer Applications*, vol. 15, pp. 31–45, January 1992.

[12] G. Bin, X. Gao, Y. Wang, B. Hong, and S. Gao, "VEP-based brain-computer interfaces: time, frequency, and code modulations [Research Frontier]," *Computational Intelligence Magazine, IEEE*, vol. 4, no. 4, pp. 22–26, 2009.

[13] S. Golomb, "Shift Register Sequences," *San Francisco, HoldenDay, ISBN 0894120484*, 1967.