

Auditory BCI Research Using Spoken Digits Stimulation and Dynamic Stopping Criterion

Ying Zhang, Lei Wang, Miaomiao Guo, Lei Qu, Huanhuan Cui, Shuo Yang

Province-Ministry Joint Key Laboratory of Electromagnetic Field and Electrical Apparatus Reliability, Hebei University of Technology, Tianjin, China

Email: murhythm@qq.com

How to cite this paper: Zhang, Y., Wang, L., Guo, M.M., Qu, L., Cui, H.H. and Yang, S. (2016) Auditory BCI Research Using Spoken Digits Stimulation and Dynamic Stopping Criterion. *J. Biomedical Science and Engineering*, 9, 71-77.
<http://dx.doi.org/10.4236/jbise.2016.910B010>

Received: July 26, 2016

Accepted: September 20, 2016

Published: September 23, 2016

Abstract

Auditory brain-computer interfaces (BCI) provide a method of non-muscular communication and control for late-stage amyotrophic lateral sclerosis (ALS) patients, who have impaired eye movements or compromised vision. In this study, random sequences of spoken digits were presented as auditory stimulation. According the protocol, the subject should pay attention to target digits and ignore non-target digits. EEG data were recorded and the components of P300 and N200 were extracted as features for pattern recognition. Fisher classifier was designed and provided likelihood estimates for the Dynamic Stopping Criterion (DSC). Dynamic data collection was controlled by a threshold of the posterior probabilities which were continually updated with each additional measurement. In addition, the experiment would be stopped and the decision was made once the probabilities were above the threshold. The results showed that this paradigm could effectively evoke the characteristic EEG, and the DSC algorithm could improve the accuracy and communication rate.

Keywords

Brain-Computer Interface, P300, N200, Dynamic Stopping Criterion

1. Introduction

Brain-computer interface (BCI) technology aims at helping patients with severe neuro-functional disabilities to communicate with outer world without using the natural output channels of the brain. This technology is generally based on features extracted from EEG recordings and the designing of the classifier. Most of these neural features are evoked during visual stimulation paradigms such as the P300 matrix speller paradigm and the steady-state evoked potential paradigm (SSVEP) [1]. However, a number of

BCI end-users in complete paralysis state, loss of gaze control when performing visual BCI paradigms, thus new BCI paradigms only relying on brain response to auditory stimulation need to be established.

Recently, auditory BCI is usually performed as a P300 oddball paradigm; that is, visual stimuli are replaced by audio stimuli such as bells, buzzing, tones and spoken words or numbers. Halder *et al.* [2] proposed an auditory three stimulus oddball paradigm in which a series of frequent standard tones and two targets randomized in sequences of up to seven stimuli played to subjects. The results showed that the average accuracy of 78.5% was achieved. Höhne and Tangermann presented a streaming paradigm for auditory spelling [3]. In this paradigm, the letters were presented in a constant auditory stream. The subject focused their attention to the target letter within this stream, the result showed that the average accuracy of 41%. Kleih *et al.* [4] proposed an auditory spelling paradigm the WIN-speller. In the WIN-speller letters were grouped by words, such as the word KLANG representing the letters A, G, K, L, and N. Thereby, the decoding step between perceiving a code and translating it to the stimuli it represents became superfluous. The average accuracy was 84%.

In this paper, we proposed a modified auditory oddball paradigm, in which the spoken digits (voices) are used to auditory stimulus. The subject need make a response to target digits and ignore non-target digits in experiment. Superposed average method was used to exact feature, fisher discriminant analysis was used to the target detection. With the increasing of trails, the EEG response to target voice could have already reached a discriminable level before all trials were presented. To speed up the decision, a stopping criterion should be set to terminate the stimulus repeating, so the Dynamic Stopping Criterion (DSC) was used as the experimental conditions to stop.

2. Methods

2.1. Subjects and Data Recording

Six healthy subjects, age 20 to 27 years, participated in the experiment. The experimental procedure was explained in detail to each subject. None of the subjects had a history of hearing loss and were all able to clearly hear and identify the different digits used in the experiment.

A standard EEG cap (Compumedics Neuroscan, USA) with 64 surface electrodes was used for recording EEG signals. EEG data were collected from Fz, FCz, Cz, CPz, Pz, Oz and sampled at 1000 Hz. The SynAmps 2 amplifier and data acquisition system (Compumedics Neuroscan, USA) was used for signal conditioning and data acquisition (band-width 0.05 - 200 Hz). All electrode impedances were kept below 5 K Ω during data recording.

2.2. Experimental Procedure

In this experiment, two mono stimulus sequences were played simultaneously. The sequence played in the left ear consisted of four spoken digits in Chinese, *i.e.*, 1, 2, 3, 4. The sequence played in the right ear consisted of four spoken digits in Chinese, *i.e.*, 5,

6, 7, 8. One target and seven non-target stimuli were delivered in each trail randomly. One trail of the sequence presented in the ear is shown in **Figure 1**. All voices were presented for 200 ms with a randomized inter stimulus interval (ISI) of 250 - 450 ms. The loudness was set to 75 dB sound pressure level (SPL). Each sequence consisted of 40 trails. The two sequences differed in the direction from which they were presented to the subject (either the right or left channel of stereo headphones) simultaneously. The task consisted of two sessions, in each of which the subject was requested to attend to a sequence in one direction and to count the target stimuli in the attended sequence. Before each task, the participant was asked a randomly selected “yes/no” question. If the subject selected “yes”, he was supposed to attend to left target (one of 1 - 4), and if the subject selected “no”, to instead attend to right target (one of 5 - 8).

2.3. Data Analysis

The raw EEG data were filtered using a 0.5 - 10 Hz digital band-pass filter. The data epochs were extracted from 200 ms before the stimulus-onset to 800 ms after the stimulus-onset. Then, all epochs were referred to the mean amplitude of the prestimulus baseline. There were 40 trials each task, with eight stimuli in each trial, 320 epochs were extracted from each task. To determine EEG responses differences between the target and non-target, grand average analysis was used to extract the features. Fisher discriminant analysis was adopted to discriminate between EEG responses for target and non-target.

The DSC algorithm is illustrated in **Figure 2**. Fisher discriminant analysis was used with this algorithm, the classifier responses for target and non-target digits were grouped. The training data consisted of 20 target digits for each of which twenty trails were measured. Thus, the non-target group would consist of classifier responses to 140 non-target digits while the target group would consist of classifier responses to 20 target digits. Kernel density estimates using a Gaussian kernel [5] were calculated for each group to estimate the probability density function of the likelihood. These participant-specific likelihoods are the only requirement for the algorithm to control data collection in real time. The method for updating the probability that a digit is the target

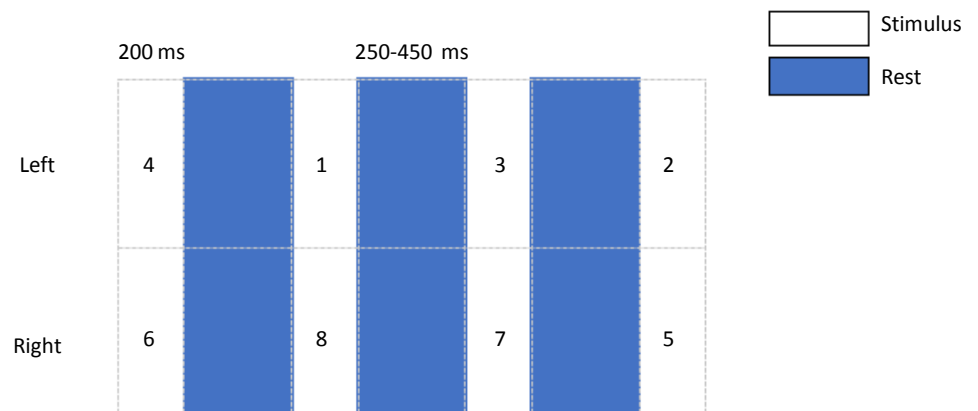


Figure 1. One trail of the sequence.

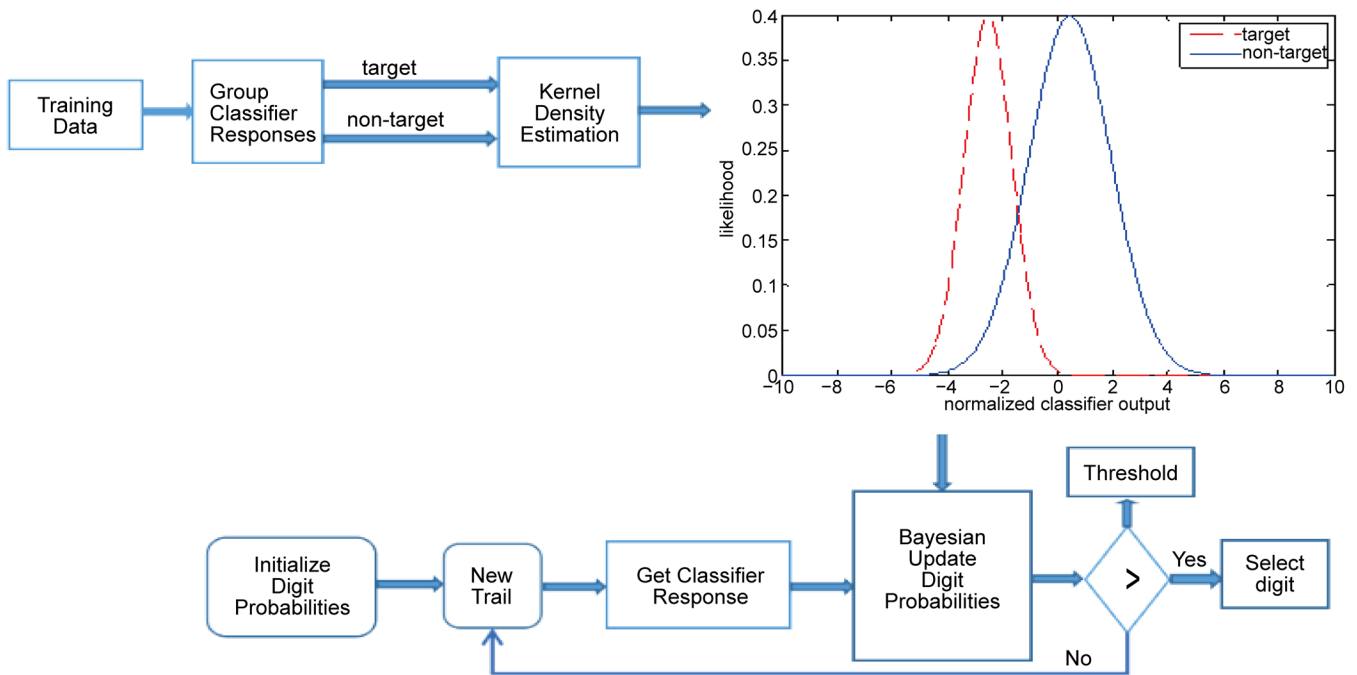


Figure 2. DSC algorithm.

digit based on previous classifier responses is based on Bayes rule.

$$p(C|X) = \frac{p(X|C)p(C)}{p(X)} \tag{1}$$

where $p(C|X)$ is the current estimate of the digit’s probability of being the target given all of the classifier responses, X , observed previously; $p(C)$ is the prior probability for the digit; $p(X|C)$ is the likelihood of the classifier responses; and $p(X)$ is the probability of the classifier responses. For an online algorithm, the posterior probabilities need to be updated after each stimulus. Hence, sequential updating of the digit probabilities was carried out using the following:

$$p(C|x_i, S_i, X) = \frac{p(C|X)p(x_i|C, S_i)}{\sum_i p(C|X)p(x_i|C, S_i)} \tag{2}$$

where $p(C|X)$ is the current estimate of the digit’s probability of being the target given all of the classifier responses observed previously; $p(x_i|C, S_i)$ is the likelihood of the current classifier response x_i , given that the digit was/was not in the currently presented set of digits (S_i); and the denominator normalizes the updated probabilities by dividing by the sum over all digits probabilities.

After each Bayesian update, the spoken digit probabilities are compared to a threshold to determine if the target digit can be selected. The threshold indicates the confidence that the correct digit has been chosen, and in this study, it was set to 90%. Once a target digit is selected, the process begins again for the next target digit by reinitializing all of the digit probabilities.

3. Results

3.1. Auditory ERP

Figure 3 shows one an example of average ERP response to target stimuli and non-target stimuli. A negative peak with latency around the 100 - 200 ms period (N200) was showed in the target stimuli response. In addition, the P300 component was elicited by target stimuli, but absent in response to non-target stimuli.

3.2. Results of DSC Algorithm

DSC algorithm was used to establish the dynamic model, and the off-line data were simulated, and the adaptive stopping criterion was obtained. **Figure 4** is the distribution of target and non-target probability density using kernel algorithm to estimate the

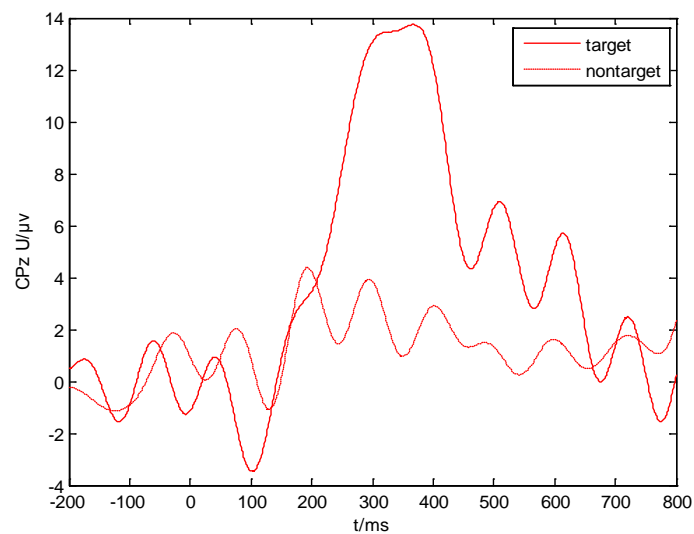


Figure 3. Grand average waveform.

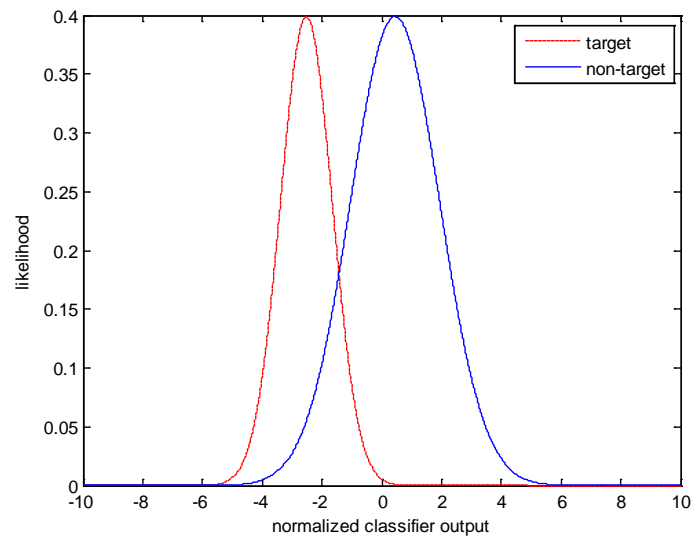


Figure 4. Likelihoods of classifier response.

output of the classifier results, in which the dotted line represents the output possibility of target digital stimulation while solid line represent the non-target. **Figure 5** shows stopping criteria probability curve, it is apparent that target stimuli are able to be selected with fewer trails. The comparison of communication rate between DSC algorithm and traditional static algorithm is presented in **Figure 6**. In this part, the communication rate was calculated using 10 trails. We found that the communication rate based on DSC condition can exceed the result of traditional static algorithm.

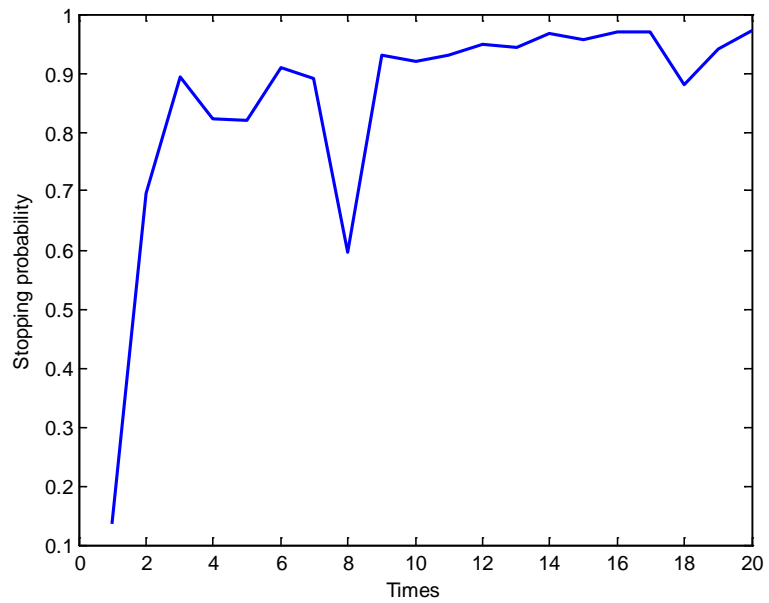


Figure 5. DSC model.

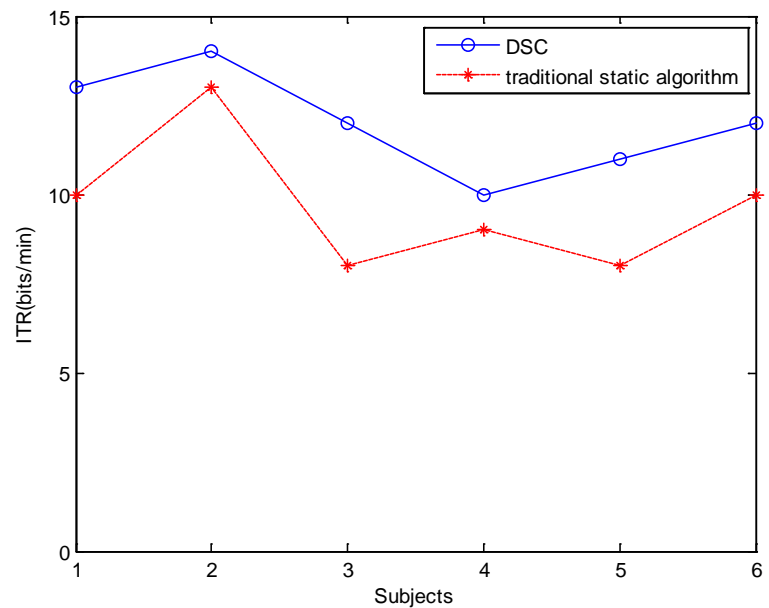


Figure 6. The comparison of communication rate between DSC algorithm and traditional static algorithm.

4. Conclusions

In this paper we presented a novel auditory BCI paradigm and acquired the EEG data from six subjects. Dynamic stopping standard based on Bayes algorithm was applied for the experiments data. In our DSC algorithm, the threshold was set as the probability of target detection and was estimated from training data, which is a reasonable prediction for final detection accuracy. The results showed that the target stimuli in the paradigm evoked obvious N200 and P300 response. Adaptive way in terms of time performance was superior to the traditional static way with higher accuracy. In a short word, the DSC algorithm of deciding the optimal number of averaging trials can improve the accuracy and communication rate of an auditory BCI.

In conclusion, the result of experiments suggests that this auditory paradigm and DSC algorithm are feasible for an auditory BCI system. Thus, it is possible to further implement an auditory BCI system with high accuracy and communication rate.

Acknowledgements

This study was supported by the National Natural Science Foundation of China, No. 31300818 and the Foundation of Hebei Province Department of Education, No. QN2016097.

References

- [1] Yin, E., Zhou, Z., Jiang, J., *et al.* (2013) A Novel Hybrid BCI Speller Based on the Incorporation of SSVEP into the P300 Paradigm. *Journal of Neural Engineering*, **10**, Article ID: 026012. <http://dx.doi.org/10.1088/1741-2560/10/2/026012>
- [2] Halder, S., Rea, M., Andreoni, R., *et al.* (2010) An Auditory Oddball Brain-Computer Interface for Binary Choices. *Clinical Neurophysiology*, **121**, 516-523. <http://dx.doi.org/10.1016/j.clinph.2009.11.087>
- [3] Höhne, J. and Tangermann, M. (2014) Towards User-Friendly Spelling with an Auditory Brain-Computer Interface: The Charstreamer Paradigm. *PloS One*, **9**, e98322. <http://dx.doi.org/10.1371/journal.pone.0098322>
- [4] Kleih, S.C., Herweg, A., Kaufmann, T., *et al.* (2015) The WIN-Speller: A New Intuitive Auditory Brain-Computer Interface Spelling Application. *Frontiers in Neuroscience*, **9**. <http://dx.doi.org/10.3389/fnins.2015.00346>
- [5] Bishop, C.M. (2006) Pattern Recognition and Machine Learning. Springer Science/Business Media, New York.

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact jbise@scirp.org