35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Sequential Selection of Window Length for Improved SSVEP-Based BCI Classification

Erik C. Johnson, James J.S. Norton, David Jun, Timothy Bretl, Douglas L. Jones

Abstract— **Brain-computer interfaces (BCI) utilizing steady-state visually evoked potentials (SSVEP) recorded by electroencephalography (EEG) have exciting potential to enable new systems for disabled individuals and novel controls for robotic and computer systems. To interact with SSVEP-based BCIs, users attend to visual stimuli modulated at predetermined frequencies. A key problem for SSVEP-based BCIs is to classify which modulation frequency the user is attending, for which there is an inherent trade-off between speed and accuracy. As SSVEP signals vary with time and stimulation frequency, a fixed-length data window does not necessarily optimize this trade-off. We propose a strategy, developed from sequential analysis, to vary the window-length used for classification. Our proposed technique adapts to the data, continuing to collect data until it is confident enough to make a classification decision. Our strategy was compared to a fixed window-length method using a simple experiment involving five frequencies presented individually to three participants. Using a canonical correlation analysis classifier to compare the proposed variable-length scheme to a standard fixed-length scheme, the variable-length approach improved the classifier information transfer rate by an average of 43%.**

I. INTRODUCTION

The direct classification of neural signals for the benefit of those with motor impairments or as an alternate input modality has intrigued researchers since first proposed more than 40 years ago [16]. Using electroencephalography (EEG), researchers have continued to develop brain computer interfaces (BCI) for communication and control, effectively translating electrical brain activity into artificial commands. One class of BCI systems, based on steady-state visually evoked potentials (SSVEP), relies on the brain's response to repetitive visual stimuli in the user's environment [13]. These stimuli, such as a flashing LED or computer screen, cause an entrainment between populations of neurons and the stimuli that can be selectively modulated through the allocation of attention [11]. Based on the neural signatures measured by EEG, this allocation of attention can be classified, effectively allowing the user to control the input of a computer system without the need for motor interaction. The BCI devices based on this paradigm have enabled text communication for disabled individuals [17], been demonstrated for use in robotic navigation tasks [1], and have been used as inputs for computer games [8].

Despite the promise demonstrated using these techniques their utility remains limited for several reasons, including: their reliability [2], ease of use, and overall system performance [12]. Considering the third of these three limitations, BCI systems are commonly compared based on their Information Transfer Rate (ITR), measured in bits/second [15]. Since ITR is a function of accuracy, latency, and the number of classes, various schemes for improving bitrate can be imagined. For instance, an easy way to improve overall information throughput is to increase the number of classes available to the user. Even if classification is relatively slow, a high ITR can be obtained from a system with 48 classes [5]. Another approach is to improve the speed and accuracy of classification.

Several classification methods currently dominate SSVEP-based BCI systems, including: Power Spectral Density Analysis (PSDA), Minimum Energy Combination (MEC) [4], and Canonical Correlation Analysis (CCA) [10]. In order to classify which of several frequencies a user is attending to, these classifiers wait for a fixed length of input EEG data before making a decision. This "window-length" is chosen *a priori* by the system designer and represents a trade-off between classification speed and accuracy.

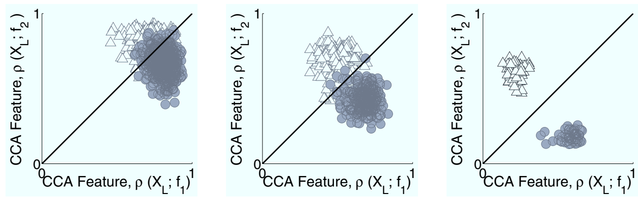
However, there is no basis for assuming the window length must be fixed. Sequential analysis [9] provides a framework that makes selecting the stopping time (i.e., choosing the window length) a part of the classification task. This methodology is commonly used in a wide range of applications including medical diagnostics and quality assurance in manufacturing. Straightforward application of the sequential probability ratio test (SPRT) [18] (a standard sequential procedure), to SSVEP classifiers is not immediately obvious due to the lack of appropriate signal models.

In this paper we propose a sequential test which is performed directly in the classifier-feature space. This test accounts for the classification rule, does not require modeling assumptions, and can handle nonlinear feature mappings. We develop our variable-length window method for the CCA classifier presented in [10], although the methodology outlined in this paper may be extended to other existing classification methods. We demonstrate, based on a comparative study with a traditional CCA algorithm, that our variable-length window method allows for classification on a short window-length when signal quality is high, and automatically waits for more data when signal quality is low. This sequential approach is shown to uniformly outperform a fixed

This work is supported by the University of Illinois Neuroengineering IGERT under NSF Grant No. 0903622, the NSF EFRI-BSBA Award No. 0938007, and the Multiscale Systems Center, a Semiconductor Research Corporation program.

E. Johnson, D. Jun, and D. Jones are with the Department of Electrical Engineering, J. Norton is with the Neuroscience Program, and T. Bretl is with the Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA {ejohns24,davidjun,dl-jones,jnorton4,tbretl} at illinois.edu

978-1-4577-0216-7/13/$26.00 ©2013 IEEE



(a) $L = 16$ (0.125 s) (b) $L = 64$ (0.5 s) (c) $L = 256$ (2 s)

Fig. 1. Plots of the CCA classifier decision rule for data of increasing window lengths. Test data with $f_1 = 6$ Hz and $f_2 = 8.57$ Hz were divided into three different lengths. As window length increases, the discriminability of the two classes increases considerably.

window-length method in terms of ITR and classification accuracy.

II. METHODOLOGY

In this section, we first introduce notation for the CCA classifier [10], which is then used to develop the proposed variable-length window method.

A. Stimulus-Frequency Classification using CCA Features

The goal of the classification algorithm proposed in [10] is to infer the input frequency from multi-channel EEG data:

1) Assume there are K possible stimulus frequencies and N EEG channels.
2) The window length is assumed to be L samples long, which is L/F_s seconds, where F_s is the EEG sampling rate.
3) For a given stimulus frequency, CCA coefficients are computed using L samples from each of the N channels. The data are represented by a matrix X of dimension $L \times N$.
4) The largest CCA coefficient, defined to be $\rho(X; f)$ [6], is used as the feature for classification:

$$f^* = \arg\max_f \ \rho(X; f), \quad f = f_1, \ldots, f_K \qquad (1)$$

where f^* is the classified frequency.

B. A Sequential Approach to a Variable-Length Window

The proposed sequential test essentially uses all of the data collected thus far to decide whether or not to continue collecting more data. To avoid requiring any additional assumptions about the data model, we develop a sequential test directly on the CCA feature space. This allows for direct comparison between the variable-length window and standard CCA classification.

The decision to continue collecting data depends on the current level of confidence as to which class the data belongs. The classification strategy in (1) for two classes f_1 and f_2 (i.e., $K = 2$) is shown graphically as the solid line in each subplot of Fig. 1. Each subplot corresponds to a different window length L, and each sample corresponds to the features extracted from an $L \times N$ block of raw EEG data. All features lying above this line are classified as f_2, and all features below it are classified as f_1. Thus, the classification boundary can be expressed as a hyperplane, h. Given the features of a data block (ρ_1 and ρ_2), the distance from the classification boundary, which expresses one's confidence, is given as the distance from the hyperplane:

$$d(\rho_1, \rho_2) = \left| h^T \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \right| = \left| [1, -1] \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \right| = |\rho_1 - \rho_2| \qquad (2)$$

As shown in the sequence of scatter plots, when the window-length L increases, the confidence for every point also increases. This behavior has been quantified in the related problem of detecting a sinusoidal signal in Gaussian noise, where the Chernoff distance between the null and alternative densities (or more generally the deflection coefficient) increases by a factor of $\sqrt{L}$ [9].

The key idea of our approach is that even when L is small, there are *some* samples that could be classified correctly with high confidence. Identifying these samples and classifying them early may help to reduce the *effective* window length, when averaged over time.

Given a current window length L and data X (of dimension $L \times N$), we propose the following sequential test:

Require: confidence threshold, τ
Output: classification result, f^*

procedure VARYWINDOWLENGTH(L, X)
  if $d(\rho(X; f_1), \rho(X; f_2)) < \tau$ **then**
    $L \leftarrow L + B$ ▷ *increase window length*
    $X \leftarrow [X; X_B]$ ▷ *augment new data*
    return `VaryWindowLength`(L, X) ▷ *repeat*
  else
    return f^* from (1) ▷ *classification task*

Here, B is defined to be a step-size, which is a basic unit of growth, and X_B is the data matrix that corresponds to the new B samples for each of the N channels. The classification task is carried out only when the minimum distance from the boundary is satisfied; until then, the window length is incrementally increased.

The threshold τ controls the trade-off between classification performance and average speed. In practice, choosing the threshold should be done *a priori*, and is exactly analogous to how window-length is chosen in a fixed-length strategy. Although the procedures are similar, the impact is quite different, as our variable-length strategy will be able to make classifications early, with minimal effect on performance. In fact, a fixed-length strategy is actually a special case of our proposed algorithm, with $\tau = 0$, and the initial L chosen to be the fixed window-length.

Finally, note that we focus on binary classification for the remainder of the paper. The generalization ($K > 2$) is a straightforward extension using the geometric hyperplane interpretation, and will be explored in a forthcoming paper.

III. EXPERIMENTAL SETUP

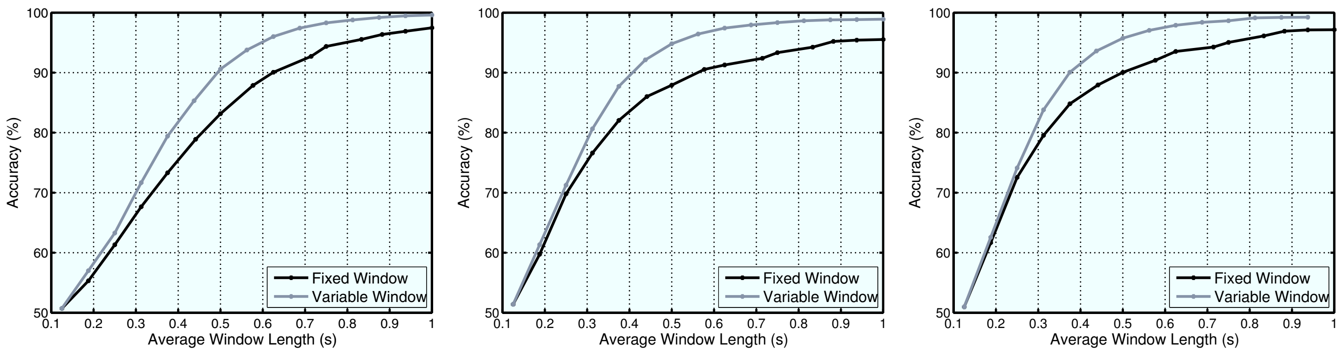
A. Subjects

Experiments were conducted at the University of Illinois BCI lab on the authors. A James Long 128 channel EEG

TABLE I

PERFORMANCE OF 3 SUBJECTS FOR FIXED-LENGTH AND VARIABLE-LENGTH WINDOW

	CCA (fixed-length window)			CCA (variable-length window)			
Participant	Accuracy (%)	AWL (s)	Max ITR (bits/s)	Accuracy (%)	AWL (s)	Max ITR (bits/s)	% Max ITR Improvement
A	94%	0.75	0.92	96%	0.62	1.24	35.3%
B	90%	0.58	0.96	94%	0.48	1.44	49.8%
C	88%	0.44	1.09	95%	0.48	1.59	45.2%
Average	91%	0.59	0.99	95%	0.53	1.42	43.4%



(a) Subject A (b) Subject B (c) Subject C

Fig. 2. Classification accuracy vs average window length for all three subjects averaged over 10 two-class comparison cases. For the fixed-length window, the performance curve is generated by changing the length of the fixed window. For the variable-length case, the performance curve is generated by altering the decision threshold. These curves were then linearly interpolated and averaged together across all 10 comparisons. For all three subjects, the variable-length window performs uniformly better, and approaches the performance of the fixed-window only for low average window lengths.

amplifier was used in conjunction with a National Instruments DAQ to digitize EEG signals at 128 Hz. The data were pass-band filtered from 1 – 30 Hz by the amplifier. EEG data were monitored during experimentation and logged by BCI2000 [14]. Participants were seated in a comfortable chair at 65 cm from a 24-inch BenQ XL2420T computer monitor. Scalp recording impedances were kept under 10kΩ from sites (PO7, PO3, PO4, PO8, O1, OZ, O2) based on the 10-5 international system [7].

B. Stimuli and Procedure

Stimuli were implemented as a script in MatLab in conjunction with the Psychophysics Toolbox [3]. The experiment consisted of six blocks of stimuli (6 Hz, 6.67 Hz, 7.5 Hz, 8.57 Hz, 10 Hz, and Null). During a block, a single stimulus of a given frequency was presented to the participant. The stimuli within each block were not randomized for this study as the emphasis was on a direct comparison between the two algorithms. Each stimulus was a square of identical size, subtending an angle 3.5° from fixation in each direction. Each block was composed of 20 trials each 15 seconds in length. At the beginning of each block, the participant was instructed to focus his attention on the center of the flickering stimuli for the entire duration of the trial. There was a three-second interval between each trial. Stimulus onset was captured with a photodiode linked directly into the DAQ.

C. Analysis Techniques

All analyses were conducted offline following each experiment in the MatLab environment. In order to quantify the difference in performance between the fixed-length window and our proposed variable-length window, the strategies were tested using a set of two-class classification problems. To form the two-class problems, each stimulus frequency was compared, one at a time, against all other frequencies. This gave a total of 10 comparisons for each of the three subjects. For each two-class problem, 20 trials of each frequency formed the testing dataset. Each two-class problem, therefore, had 10 minutes of testing data. A two-class CCA classifier was applied to the testing dataset, using both a fixed-length and variable-length window strategy.

D. Parameter Selection

Performance for the fixed-length window strategy was tuned by modifying the window length. The window length varied from $1/8$ seconds to 1 second with $1/16$ second steps. Changing the length of the fixed-length window trades off between decision speed and decision accuracy. For the variable-length strategy, the minimum block length was set at $1/8$ second. To tune performance of the variable-length strategy, the threshold τ was varied from 0 to 0.3 in steps of 0.01. Varying the threshold trades off between classification speed and accuracy.

IV. RESULTS

For both fixed-length and variable-length strategies, the percent accuracy, average window-length (AWL), and ITR was calculated and averaged over all 10 two-class comparisons. The maximum ITR for each subject is reported in Table I for both fixed and variable-length strategies.

The variable-length approach increases ITR by an average of 43% over all three subjects. For our subjects, using a variable-length strategy is an effective way to improve the performance of a CCA classifer.

In addition to the parameter configuration that maximizes ITR, it is possible to trace out the performance trade-off between classification accuracy and average window length for the two strategies. The results are summarized in Fig. 2. For all three subjects, the classification performance curves for the variable-length window exceeds the curve for the fixed-length window. As the threshold for the variable-length window is lowered, it approaches the performance for the fixed-length window.

V. Discussion

The maximum ITR achieved by each of the three subjects in the variable-length window exceeded the performance of the fixed-length window. This is very encouraging, and shows uniform improvement of our variable-length approach. Performance numbers, however, are derived from idealized comparisons of ITR and are not directly comparable to performance numbers from real-time BCI systems. The relative improvement of the variable-length approach does suggest that incorporating this strategy will improve performance.

Using this simple experimental data and CCA classification, the comparisons of fixed-length and variable-length windows validate the intuition for applying sequential analysis. Because the quality of the data varies with time, a variable-length window can exceed the performance of a fixed-length window. We hypothesize that the variable-length strategy can be applied to other SSVEP classifiers, such as PSDA and MEC [4], provided the performance of the classifier improves as a function of the length of data used to classify.

Although the results in this study only consider the two-class case, the variable-length strategy can be extended to the multiple-class case. For CCA, this would involve finding the CCA correlation for each frequency of interest. These features would form an n-dimensional hypercube, with hyperplane decision boundaries.

Finally, this study does not apply any channel selection or denoising techniques. Again, this is because these results demonstrate the relative improvement of applying a variable-length window in place of a fixed-window, not an absolute performance metric.

VI. Conclusion and Future Work

Since SSVEP signals vary with time, conditions, and stimulation frequencies, fixed-length windows are not necessarily optimal. This work proposed a variable-length window method for classification using CCA for SSVEP-based BCI. Our intuition about performance varying over time is consistent with the obtained results. In particular, a variable window-length strategy is shown to be uniformly better than a fixed window-length strategy, resulting in an average ITR improvement of 43%. As demonstrated, our proposed approach does not require any additional assumptions or signal models relative to existing CCA-based classifiers.

One implication of the achieved performance improvement is that faster, more accurate classification may improve the overall usability of SSVEP-based BCI systems. This may be particularly important for long-term applications, where attention and signal quality is expected to vary greatly due to effects such as fatigue and variable recording conditions.

As our approach naturally lends itself to extensions, further studies will consider multiple classes and other classification algorithms. Although current results demonstrate the relative improvement of the variable-length strategy over the fixed-length strategy, they do not yet demonstrate the performance of a real-time BCI system; future work will explore the efficacy of this approach for online BCI tasks.

VII. Acknowledgment

The authors would like to thank Abdullah Akc̃e for his constructive suggestions.

References

[1] A. Akce, J. Norton, and T. Bretl. An EEG-based brain-computer interface for indoor navigation along human-like paths. In *2012 International Conference on Intelligent Robots and Systems*, 2012.

[2] B.Z. Allison, E.W. Wolpaw, and J.R. Wolpaw. Brain–computer interface systems: progress and prospects. *Expert Review of Medical Devices*, 4(4):463–474, 2007.

[3] D.H. Brainard. The psychophysics toolbox. *Spatial Vision*, 1997.

[4] O. Friman, I. Volosyak, and A. Graser. Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces. *IEEE Transactions on Biomedical Engineering*, 54(4):742–750, 2007.

[5] X. Gao, D. Xu, M. Cheng, and S. Gao. A BCI-based environmental controller for the motion-disabled. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):137–140, 2003.

[6] H. Hotelling. Relations between two sets of variables. *Biometrika*, 28(3):321–377, 1936.

[7] V. Jurcak, D. Tsuzuki, and I. Dan. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. *Neuroimage*, 34(4):1600–1611, 2007.

[8] E.C. Lalor, S.P. Kelly, C. Finucane, R. Burke, R. Smith, R.B. Reilly, and G. Mcdarby. Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment. *EURASIP Journal on Applied Signal Processing*, 2005.

[9] B.C Levy. *Principles of Signal Detection and Parameter Estimation*. Springer, New York, NY, 2008.

[10] Z. Lin, C. Zhang, W. Wu, and X. Gao. Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. *IEEE Transactions on Biomedical Engineering*, 53(12):2610–2614, 2006.

[11] ST Morgan, JC Hansen, and SA Hillyard. Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences*, 1996.

[12] F. Popescu, B. Blankertz, and K.R. Müller. Computational challenges for noninvasive brain computer interfaces. *IEEE Intelligent Sys.*, 2008.

[13] D. Regan. Steady-state evoked potentials. *Journal of the Optical Society of America*, 67(11):1475–1489, 1977.

[14] G. Schalk, D.J. McFarland, T. Hinterberger, N. Birbaumer, and J.R. Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Trans. Biomedical Engineering*, 51(6):1034–1043, 2004.

[15] G. Schalk, J.R. Wolpaw, D.J. McFarland, and G. Pfurtscheller. EEG-based communication: presence of an error potential. *Clinical Neurophysiology*, 111(12):2138–2144, 2000.

[16] J.J. Vidal. Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, 2(1):157–180, 1973.

[17] I. Volosyak, H. Cecotti, D. Valbuena, and A. Graser. Evaluation of the bremen SSVEP based BCI in real world conditions. In *IEEE Int. Conf. on Rehabilitation Robotics*, pages 322–331, 2009.

[18] A. Wald. Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, pages 117–186, 1945.