Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs

Zhonglin Lin*, Changshui Zhang, *Member, IEEE*, Wei Wu, and Xiaorong Gao

***Abstract*—Canonical correlation analysis (CCA) is applied to analyze the frequency components of steady-state visual evoked potentials (SSVEP) in electroencephalogram (EEG). The essence of this method is to extract a narrowband frequency component of SSVEP in EEG. A recognition approach is proposed based on the extracted frequency features for an SSVEP-based brain computer interface (BCI). Recognition Results of the approach were higher than those using a widely used fast Fourier transform (FFT)-based spectrum estimation method.**

***Index Terms*—Brain computer interface, canonical correlation analysis, electroencephalogram, steady-state visual evoked potentials.**

I. Introduction

STEADY-STATE visual evoked potentials (SSVEP) can be defined as periodic evoked potentials induced by rapidly repetitive visual stimulation, typically at frequencies greater than 6 Hz. The SSVEP is composed of a number of discrete frequency components. The frequency range associated with the SSVEPs normally comprise the fundamental frequency of the visual stimulus as well as its harmonics. SSVEPs are often recorded noninvasively by electroencephalography (EEG).

Several SSVEP-based brain computer interface (BCI) systems have already been proposed [1]–[3]. Among these, the system proposed in [3] attained a high information transfer rate, reaching approximately 60–90 bits/min. The main part of the system included an LED panel. Each LED flickered at a specific frequency representing a specific command. For example, an LED flickering at a frequency X Hz would be associated with the command "turn off music." A BCI user would be able to express this command (i.e., "turn off the music") simply by gazing at the corresponding LED. An SSVEP with a fundamental frequency similar to that of the flickering LED would be induced and observed in the subject's EEG. The user's intended command can, therefore, be determined by analyzing the EEG signals, extracting the frequency information associated with the SSVEP. This is called "frequency-coded selection by SSVEP."

The most widely used frequency detection method in SSVEP-based BCIs is *power spectral density based analysis (PSDA)*. Power spectral density is estimated from the user's EEG signal within a time window and its peak subsequently detected. The frequency corresponding to this peak is taken as the visual stimulus frequency. The periodogram is a nonparametric power spectrum estimation method that can be calculated directly from the discrete Fourier transform (DFT) [4]. Therefore, the fast Fourier transform (FFT) can be utilized in its calculation, and the computational cost is low. Compared to parametric power spectrum estimation methods, the periodogram does not need a prior order selection step and can easily be implemented. Moreover, when data are contaminated by noise, the periodogram approach has been shown to be more robust than parametric estimation methods [5].

Although FFT-based PSDA works well in BCI systems, it has its drawbacks. PSDA might still be sensitive to noise if the signal to be analyzed is from a single channel (or a bipolar montage). Array signal processing, such as canonical correlation analysis (CCA) using channel covariance information, may improve the signal-to-noise ratio (SNR). We propose a frequency recognition approach based on CCA, which makes use of signals from multiple channels.

CCA is a multivariable statistical method used when there are two sets of data, which may have some underlying correlation. CCA extends ordinary correlation to two sets of variables [6], [7]. First, CCA finds a pair of linear combinations, called canonical variables, for two sets, such that the correlation between the two canonical variables is maximized. Then it finds a second pair, which is uncorrelated with the first pair of canonical variables but has a next highest correlation. The process of constructing canonical variables continues until the number of pairs of canonical variables equals the number of variables in the smaller set. The coefficients describe the correlation relation of the two sets. Note that although CCA generates multiple correlation coefficients, in this paper we only consider the largest coefficient, which has the best description capacity. Ordinary correlation studies the relationship between two variables, whereas CCA studies the relationship between two sets of variables, which is the case for some real-life problems [8], [9]. The method proposed in this paper focuses on the relationship between the stimulus signals and EEG signals from multiple channels in a local area.

II. Methods

A. Subjects and Experimental Settings

SSVEP-based BCI systems were designed and offline analyses performed. The data were recorded in our lab experiments and the hardware settings were described in [10]. A virtual keyboard displayed on a CRT screen functioned as a visual stimulator. Each key flickered at a certain frequency

Manuscript received April 18, 2006; revised September 24, 2006. This work was supported in part by the Chinese Ministry of Education under Key Project 1041185 and in part by the National Science Foundation of China (NSFC) under RGC Joint Research Grant 60318001. *Asterisk indicates corresponding author.*

*Z. Lin is with the Department of Automation, Tsinghua University, Room 3–120, FIT Building, Beijing 100084, China (e-mail: linz102@mails.tsinghua.edu.cn).

C. Zhang is with the Department of Automation, Tsinghua University, Beijing 100084, China.

W. Wu and X. Gao are with the Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, China.

Digital Object Identifier 10.1109/TBME.2006.886577

and constituted a frequency-coded matrix. The corresponding SSVEP would emerge when the subject gazed at one of the keys. The EEG signals were recorded using a BioSemi ActiveTwo system (BioSemi, Amsterdam, The Netherlands). Electrode placement followed the 64-channel extended international 10/20 system, and the sampling rate was 2048 Hz. The reference channel was Cz. Downsampling to 256 Hz was finally performed. The offline systems each contained 9 stimulus frequencies (i.e., K = 9) of 27, 29, 31, …, 43 Hz. Eleven volunteers took part in the study. Each subject carried out 9 runs corresponding to the 9 frequencies. Each run lasted for 60 s and the signals were recorded. Between runs, subjects could relax in order to eliminate the affect of visual fatigue. In our offline analyses, the data were band-pass filtered between 22–48 Hz (infinite impulse response filtering with order 10).

B. EEG Frequency Component Analysis Based on CCA

CCA works on two sets of variables. In our method, variables in one set are the signals, $x(t)$, recorded from several channels within a local region and the second set is from stimulus signals. Any periodic signal can be decomposed into a set of Fourier series. For example, a square-wave periodic stimulus signal, $y(t)$, at a certain frequency f can be decomposed into the Fourier series of its harmonics $(\sin(2\pi ft), \cos(2\pi ft), \sin(4\pi ft), \ldots)$

$$y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \\ y_4(t) \\ y_5(t) \\ y_6(t) \end{pmatrix} = \begin{pmatrix} \sin(2\pi ft) \\ \cos(2\pi ft) \\ \sin(4\pi ft) \\ \cos(4\pi ft) \\ \sin(6\pi ft) \\ \cos(6\pi ft) \end{pmatrix}, \quad t = \frac{1}{S}, \frac{2}{S}, \ldots, \frac{T}{S} \tag{1}$$

where f is the base frequency, T is the number of sampling points, and S is the sampling rate. Fig. 1 illustrates the use of CCA in EEG signal analysis. Due to the fact that brain dynamics perform as a low-pass filter [11], [12] high harmonic components in a square wave may be filtered. As a result, we use the 6 harmonics as in (1) in the analysis. Our working hypothesis is that the activity of the electrical source responsible for SSVEP is the output of a linear system with the stimulus signal as the input. Consequently, it follows that SSVEP contains the same frequency as the stimulus signal.

In this study, we employed a feature extraction method based on the CCA method for SSVEP detection and recognition. In this approach, EEG signals from multiple channels are used to calculate the CCA coefficients with all stimulus frequencies in the system. The frequency with the largest coefficient is the one of SSVEP.

C. Recognition Strategy

The core problem in SSVEP-based BCI systems is to detect the frequency of the SSVEP component in the subject's EEG. Suppose there are K stimulus frequencies $f_1, f_2, \ldots, f_K$ and that the analyzed signal has been acquired from N channels within an L s window. Our strategy for recognition is as follows. The stimulus frequency f_s, satisfies

$$f_s = \max_f \rho(f), \quad f = f_1, f_2, \ldots, f_K \tag{2}$$

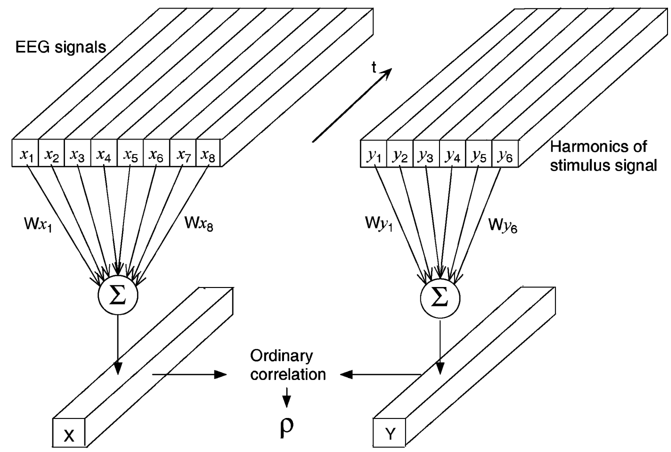


Fig. 1. An illustration for usage of CCA in EEG signal analysis. $x_1, \ldots, x_8$ are signals from 8 EEG channels and $y_1, \ldots, y_6$ are Fourier series of a given frequency period signal. The CCA finds the linear combination coefficients $w_{x_1}, \ldots, w_{x_8}$ and $w_{y_1}, \ldots, w_{y_6}$, which gives the largest correlation between X and Y. For brief statement, we omit the variable t. This figure uses [13] for reference.

where $\rho(f)$ is the CCA coefficient of x_{LN} and y. The setting of y is indicated in (1).

D. Channel Selection

Pertinent channels for the user's SSVEP should first be determined. SSVEP is a localized potential [14]. Taking too large an area for analysis makes it difficult to find the features (i.e., the frequency of the periodic wave) using CCA becauseof the introduction of less relevant EEG features. Analysis is, therefore, carried out on a scalp *patch*. Scientists have gained some knowledge of the scalp distribution of SSVEP. From this knowledge, we can determine which channels would provide the best signals for each of the subjects. However, close attention should be paid to the inherent differences among the subjects. We have, consequently, proposed a method for channel selection, which is also based on CCA.

For each channel, we take its M nearest neighboring channels to construct a local area with $M + 1$ channels. The signals from these channels are taken to compute the CCA coefficient with the stimulus signal and the coefficient is designated to this channel. The importance of the channel in the EEG can be determined by comparing the coefficients. After averaging the CCA coefficients across all K stimulus frequencies in the BCI system, the N channels with the largest coefficients are selected to construct the local scalp patch. The signals from these channels are $x_1, x_2, \ldots, x_N$ in Fig. 1. In our analysis, we empirically set the two parameters as $N = 8$ and $M = 5$.

E. Contrast Method

For the purpose of comparison, we also implemented the PSDA method and utilized the result as a to contrast to that of our proposed CCA approach. First, the channels for PSDA were selected by the method proposed in [15]. The method selected an optimal bipolar lead with a high SNR. Then the power spectrum was estimated using the differential signal of this bipolar lead. We set the number of points to 256 for the FFT to ensure that the resolution of the power spectral density is 1 Hz. We could do 256-point FFT when L was 1 s (since the sampling rate is 256 Hz). A Hamming window of similar length to each segment was used. Zero-padding was performed when L was shorter than 1 s and truncation was performed when L was longer than 1 s. When L was equal to or longer than 2 s, two continuous segments of the 256-point signal were used to estimate the power spectral density and their average was used for peak detection. When performing recognition, we took into account the fact that all systems had only odd stimulus frequencies. Let $P(f)$ be the power density value at frequency f, and f_{peak} be the frequency corresponding to the peak of the spectrum, then the determined frequency of the PSDA approach is

$$f_s = \begin{cases} f_{\text{peak}}, & \text{when } f_{\text{peak}} \text{ is odd} \\ \max\limits_{f} P(f),\ f = f_{\text{peak}} \pm 1, & \text{when } f_{\text{peak}} \text{ is even.} \end{cases} \tag{3}$$

F. Evaluation Methods

A cross-validation design was employed to evaluate the two methods. The sample data set was divided into the training set for channel selection and the test set for testing the performance of frequency recognition. For every system, 30-s-long signals randomly cut from all stimulus frequencies constituted the training set. The remaining 30-s-long signals made up the test set. From those test sets, 900 segments of the signal of length, L s, were randomly selected (100 for an individual frequency) and frequency recognition was performed. The analysis was repeated 10 times. We studied the influence of L on recognition accuracy. The accuracy was defined as the percentage of samples (segments of signal) whose stimulus frequencies were judged correctly. Let f_s be the frequency determined by the recognition algorithm and f_{stimulus} be the real stimulus frequency of a sample. Then the accuracy can be formulated as

$$\frac{\text{acc} = \text{number of samples}(f_s = f_{\text{stimulus}})}{\text{number of samples} \times 100\%} \tag{4}$$

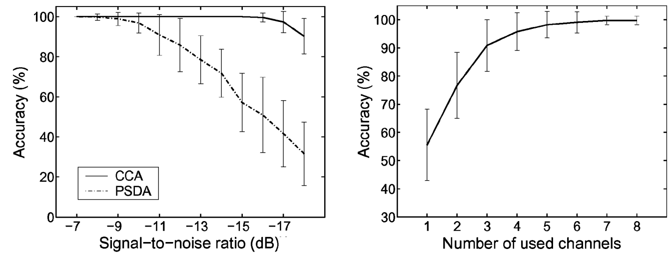
where $\text{number of samples} = 900$.

G. Simulation

We studied the anti-noise capability of the two approaches through simulation (Simulation I). First, we generated 8 sinusoidal waveforms to simulate the SSVEPs of 8 channels at each of the 9 stimulus frequencies described previously in our experiment. The sampling rate was 256 Hz and each signal lasted for 1 s. Then we added Gaussian white noises to the sinusoidal signals and achieved noise contaminated signals. Finally, we performed frequency recognition using the two approaches. The procedure was repeated 50 times, adjusting the SNR and observed the influence of SNR on the accuracy of recognition. The SNR is defined as

$$\text{SNR} = 10\log_{10}\frac{P_{\text{signal}}}{P_{\text{noise}}} = 10\log_{10}\frac{(A/\sqrt{2})^2}{\sigma^2} \tag{5}$$

where P denotes power, A is the amplitude of the sine wave, and σ^2 is the variance of the noise. Fig. 2(a) shows the recognition accuracy of the two approaches at various SNR. When SNR gets lower, the recognition accuracy of the PSDA approach falls rapidly, whereas the performance of the CCA approach remains quite good.



(a) (b)

Fig. 2. (a) The influence of noise on the two recognition approach and (b) the influence of number of channel on the CCA approach, the SNR is −15 dB were studied by simulation. The error bars represent standard deviations. (a) Shows that the CCA approach is more robust than the PSDA approach. (b) Shows that using more channels for the CCA approach may improve recognition accuracy. (a) Simulation I; (b) Simulation II.

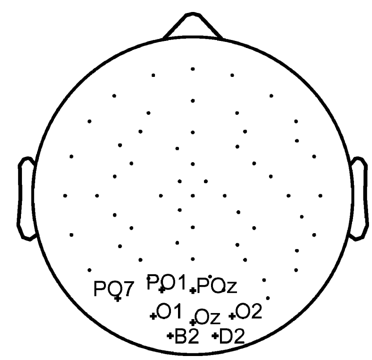


Fig. 3. The channels selected for subject CSQ in one of the 10 runs. The dots represent the locations of the electrodes on the scalp. The selected channels are labeled with their names and plus signs. They are mainly located in the occipital region. For this subject, the significant region is shifted to the left, which illustrates the specificity in subject.

We also studied the effect of the number of used channels in the CCA approach (Simulation II). Fig. 2(b) illustrates the variation in the recognition accuracy with the number of used channels. We can see from this plot that using more channels improves the recognition accuracy. However, in the real situation, not so many channels' signals contain SSVEPs. The better way is to select an appropriate number of channels, using the method in Section II-D.

III. Results

As for each subject, the channels selected by our methods differed slightly across runs in cross-validation. They mainly located in the occipital region. For example, the channels selected for subject CSQ were $B2$, Oz, $O2$, POz $D2$, $O1$, $PO7$, and $PO1$ in one of the 10 runs (Fig. 3). These results are consistent with the fact that SSVEPs were generated in the occipital region. We can see from the figure that the significant region is shifted to the left for this subject, which illustrates subject specificity.

Fig. 4 presents the overall comparisons between the two recognition methods applied for these subjects. The figure indicates that our method attains higher recognition accuracy at various time window lengths for most subjects (except subject NTS). The relatively smaller standard deviations of the CCA approach suggest that it is more robust.

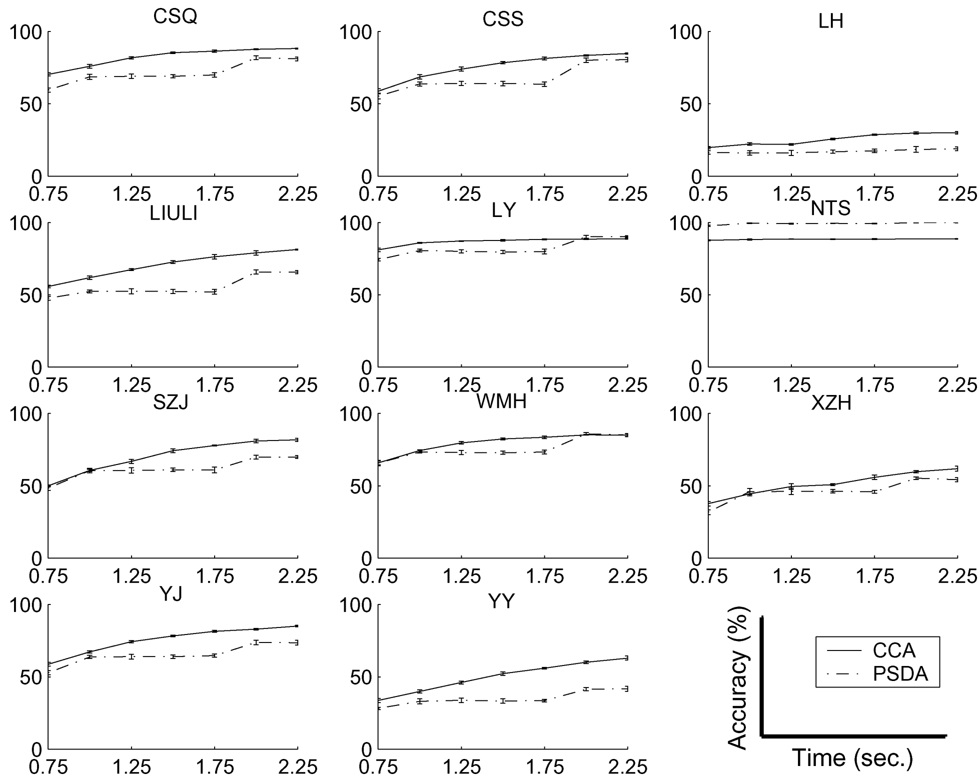


Fig. 4. Accuracy and standard deviation with respect to time window length for various subjects using the two methods. Solid lines and dashed lines represent the results of the CCA and PSDA approaches, respectively. The figure indicates that the CCA approach can get higher recognition accuracy than the PSDA approach at various time window lengths for most subjects (except subject NTS).

Fig. 5 presents the paired t-test significance test results. Black bars represent the average accuracy of eleven subjects for the CCA approach and white for the PSDA approach. The asterisk marks significance ($*: p < 0.05$; $** : p < 0.01$). The result shows that the CCA approach is significantly better than the PSDA approach at most time window lengths.

IV. Discussion and Conclusion

The fact that signals from multiple channels were used may have contributed to the much improved result associated with the CCA approach. The use of multiple channels creates greater robustness against noise. As for subject NTS, we speculate that there might be two reasons that led to the unusual results. First, the used signals in the PSDA approach had a very high SNR. Second, the area that generated the SSVEP was very small. The use of signals within a broader area in the CCA approach can introduce more noise and negatively impact recognition accuracy.

The base frequencies of the stimulus signals are relatively high in our experiment (at least 27 Hz). So, the harmonics of the SSVEPs fall in a much higher frequency band and may not be noticeable [14]. As a result, we did not intend to use the harmonics and bandpass filtered the SSVEPs. However, in those systems that reserve higher harmonics [16], the CCA approach is believed to benefit more from using higher harmonics other than the anti-noise ability. The reason is that with the chosen bases in (1), CCA can take into account the relative amplitude and phase of each frequency component [13].

In this paper, we demonstrated offline analysis. It is possible to use our approach for online BCI. There is no complex pre-

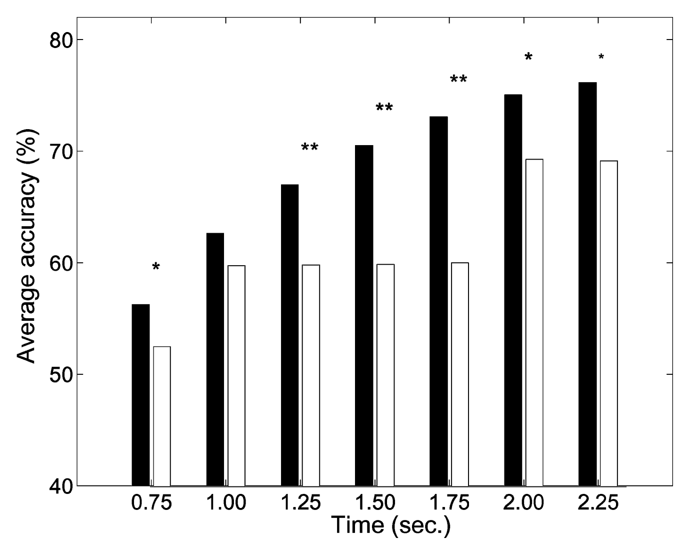


Fig. 5. Average accuracy across subjects and paired t-test result. Black and white bars represent averages of the CCA approach and the PSDA approach, respectively. The significance of paired t-test are marked by asterisks, where $*$ stands for $p < 0.05$, and $**$ stands for $p < 0.01$. The result shows that the CCA approach is significantly better than the PSDA approach at most time window lengths.

processing work other than a step for channel selection. The experiments show that the cost of this computation is negligible. Computations can be performed in 0.016 s in Matlab 6.5 environment on a PC with 1.73 GHz CPU configured for 8 channels, 2 s time window, and 9 candidate frequencies, which is sufficiently fast for practical online use. With improvements in the recognition algorithm, the BCI users may communicate faster and at the same time avoid fatigue because the eyes are subjected to the stimulus for shorter periods of time.

There have been number of feature selection methods proposed and utilized in BCI research stemming from *recursive feature elimination* (RFE) or genetic algorithms [17]–[21]. Our CCA approach may not be better than these in the sense of recognition accuracy. However, its use may benefit from easy implementation and understanding.

There are some problems with our approach. First, we only use the largest coefficient of CCA because it conveys the most information. This is true under ideal conditions. However, real EEG signals may be contaminated by noise or may have discontinuous phase transitions. These factors cause information to spread across more than one coefficient. So it might be reasonable to exploit additional coefficients. In our future work, we will attempt to study how to efficiently use more than one coefficient (e.g., use the sum of first three largest coefficients or the product of all coefficients). Second, the underlying assumption of our work is the linearity of the user's visual pathway of the brain. Nevertheless, there is considerable evidence that the human visual pathway contains multiple neural mechanisms, each of which is preferentially sensitive to one abstract feature of the visual input, and which achieve this preferential sensitivity by nonlinear processing [14]. We intend to adapt the proposed method to take into account this nonlinear effect in our future study. Third, we plan to explore how to use the projections in CCA subspaces other than only the coefficients.

This paper introduces CCA to EEG analysis for the first time. We employ CCA to extract frequency features in EEG, as well as to select channels for analysis. Then a recognition strategy for SSVEP-based BCI is proposed. Experimental results on multiple subjects showed that our approach achieved higher recognition accuracy than the widely used PSDA approach. It may be a good approach for frequency recognition in SSVEP-based BCI.

ACKNOWLEDGMENT

The authors would like to thank Y. Wang, G. A. Justin, and Dr. O. Bai for their advice, and all volunteers for taking part in the experiment.

REFERENCES

[1] G. L. Calhoun and G. R. McMillan, "EEG-based control for human-computer interaction," in *Proc. 3rd Annu. Symp. Human Interaction with Complex Systems*, Los Alamitos, CA, 1996, pp. 4–9.

[2] M. Middendorf, G. R. McMillan, G. Calhoun, and K. Jones, "Brain computer interfaces based on the steady-state visual-evoked response," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 211–214, Jun. 2000.

[3] X. R. Gao, D. Xu, M. Cheng, and S. K. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 137–140, Jun. 2003.

[4] S. L. Marple, *Digital Spectral Analysis With Applications*. Upper Saddle River, NJ: Prentice-Hall, 1987.

[5] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Aadaptive Filtering and Array Processing*. New York: McGraw-Hill, 2000.

[6] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.

[7] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 2nd ed. New York: Wiley, 1984.

[8] H. von Storch and F. Zwiers, *Statistical Analysis in Climate Research*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[9] T. Harmony, G. Hinojosa, E. Marosi, J. Becker, M. Rodriguez, A. Reyes, and C. Rocha, "Correlation between EEG spectral parameters and an educational evaluation," *Int. J. Neurosci.*, vol. 54, pp. 147–155, 1990.

[10] M. Cheng, X. R. Gao, and S. K. Gao, "Design and implementation of a brain-computer interface with high transfer rates," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 10, pp. 1181–1186, Oct. 2002.

[11] P. L. Nunez, *Electric Fields of the Brain: The Neurophysics of EEG*. New York: Oxford, 1981.

[12] C. Bédard, H. Kröger, and A. Destexhe, "Modeling extracellular field potentials and the frequency-filtering properties of extracellular space," *Biophys. J.*, vol. 86, no. 3, pp. 1829–1842, 2004.

[13] O. Friman, J. Cedefamn, P. Lundberg, M. Borga, and H. Knutsson, "Detection of neural activity in functional MRI using canonical correlation analysis," *Magn. Reson. Med.*, vol. 45, no. 2, pp. 323–330, Feb. 2001.

[14] D. Regan, *Human Brain Eelectrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. Amsterdam, The Netherlands: Elsevier, 1989.

[15] Y. J. Wang, R. P. Wang, X. R. Gao, and S. K. Gao, "Brain-computer interface based on the high-frequency steady-state visual evoked potential," in *Proc. 2005 1st Int. Conf. Neural Interface and Control*, Wuhan, China, 2005, pp. 37–39.

[16] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, "Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components," *J. Neural Eng.*, vol. 2, pp. 123–130, 2005.

[17] T. Lal, M. Schröder, T. Hinterberger, J. Weston, M. Bogdan, N. Birbaumer, and B. Schölkopf, "Support vector channel selection in BCI," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1003–1010, Jun. 2004.

[18] L. Citi, R. Poli, C. Cinel, and F. Sepulveda, "Feature selection and classification in brain computer interfaces by a genetic algorithm," in *Proc. 2004 Genetic and Evolutionary Computation Conf.*, Seattle, WA, 2004, [Online]. Available: http://www.cs.bham.ac.uk/ wbl/biblio/gecco2004/LBP073.pdf.

[19] M. Schröder, M. Bogdan, T. Hinterberger, and N. Birbaumer, "Automated EEG feature selection for brain computer interfaces," in *Proc. 1st Int. IEEE-EMBS Conf. Neural Engineering*, Capri, Italy, 2003, pp. 626–629.

[20] B. Graimann, J. E. Huggins, S. P. Levine, and G. Pfurtscheller, "Detection of ERP and ERD/ERS patterns in single ECoG channels," in *Proc. 1st Int. IEEE-EMBS Conf. Neural Engineering*, Capri, Italy, 2003, pp. 614–617.

[21] A. Rakotomamonjy, "Variable selection using SVM-based criteria," *J. Mach. Learn. Res.*, vol. 3, pp. 1357–1370, 2003.