

Recognition of Frequencies of Short-Time SSVEP Signals Utilizing an SSCCA-Based Spatio-Spectral Feature Fusion Framework

Saif Bashar

Department of CSE

*Bangamata Sheikh Fojilatunnesa Mujib
Science and Technology University*

Jamalpur, Bangladesh

Email: saifbashar2021@gmail.com

Samia Nasir Nira

Department of CSE

*Bangamata Sheikh Fojilatunnesa Mujib
Science and Technology University*

Jamalpur, Bangladesh

Email: samianira9842@gmail.com

Shabbir Mahmood

Department of CSE

*Pabna University of
Science and Technology*

Pabna, Bangladesh

Email: shabbir.cse@pust.ac.bd

Md. Humaun Kabir

Department of CSE

*Bangamata Sheikh Fojilatunnesa Mujib
Science and Technology University*

Jamalpur, Bangladesh

Email: humaun@bsfmstu.ac.bd

Sujit Roy

Department of CSE

*Bangamata Sheikh Fojilatunnesa Mujib
Science and Technology University*

Jamalpur, Bangladesh

Email: sujit@bsfmstu.ac.bd

Iffat Farhana

Department of CSE

*Varendra University
Rajshahi, Bangladesh*

Email: iffat@vu.edu.bd

Abstract—A brain-computer interface (BCI) facilitates direct communication between the brain and external equipment through EEG, which is preferred for its superior temporal resolution. Among EEG techniques, the steady-state visual evoked potential (SSVEP) is favored due to its robust signal-to-noise ratio, minimal training demands, and elevated information transmission rate. Frequency detection in SSVEP-based brain-computer interfaces commonly employs canonical correlation analysis (CCA). SSCCA (spatio-spectral canonical correlation analysis) augments CCA by refining spatial filtering. This paper presents a multistage feature fusion methodology for short-duration SSVEP frequency identification, employing SSCCA with template signals derived via leave-one-out cross-validation (LOOCV). A filterbank generates bandpass filters for stimulus frequencies and their harmonics, whereas SSCCA calculates correlation coefficients between subbands and templates. Two phases of non-linear weighting amalgamate these coefficients to discern the target stimulus. This multistage methodology surpasses traditional techniques, attaining an accuracy of 94.5%.

Index Terms—Brain-Computer Interface (BCI), Steady-State Visual Evoked Potential (SSVEP), Spatio-Spatial Canonical Correlation Analysis (SSCCA)

I. INTRODUCTION

A Brain-Computer Interface (BCI) transforms neurophysiological data into digital commands, facilitating communication with external devices without the involvement of peripheral nerves or muscles, hence providing opportunities for individuals with significant motor impairments [1]. Electroencephalography (EEG) is extensively utilized in brain-computer interfaces (BCIs) owing to its mobility, cost-effectiveness, and superior temporal resolution. SSVEP-based BCIs are favored due to their elevated signal-to-noise ratio, minimal training demands, and robust information transfer rate, utilizing vi-

sual stimuli at distinct frequencies identified by frequency recognition algorithms. Although CCA is frequently employed for frequency recognition, its accuracy limits and computational expenses have prompted the advancement of enhanced methodologies [2].

II. PROBLEM STATEMENT

CCA-based algorithms are efficient for SSVEP-based BCIs but encounter constraints such as inferior accuracy and elevated computing expenses due to dependence on predetermined reference signals devoid of EEG characteristics [3]. Advanced techniques such as MsetCCA and MCM enhance reference signals but encounter difficulties with noise extraction and insufficient utilization of SSVEP harmonics [4]. Recent methodologies such as FBMSI and SSCCA enhance accuracy and decrease computational expenses, although frequently eliminate essential discriminative information, hence affecting performance [5].

III. LITERATURE REVIEW

CCA, as presented by Lin et al. [2], use sinusoidal reference signals for frequency identification; however, it necessitates calibration and does not incorporate actual EEG characteristics. MsetCCA boosts reference signals utilizing training data [3], whereas MCM improves accuracy by diminishing noise and amplifying correlations [4]. FBMSI enhances harmonic exploitation for improved SSVEP recognition [6], while SSCCA simplifies analysis through a singular projection vector, hence augmenting efficiency and accuracy [7]. Nevertheless, conventional CCA and CORRCA preserve solely the maximum correlation coefficient, resulting in information loss [8].

To resolve this, filter bank analysis and feature fusion frameworks amalgamate subband features for enhanced frequency recognition [9], hence augmenting accuracy and information transfer rate for SSVEP-based brain-computer interfaces.

IV. DATASET DESCRIPTION

The publicly accessible dataset from the Swartz Center for Computational Neuroscience at UC San Diego derives from an offline SSVEP-based BCI experiment with 10 individuals (9 males, 1 female; average age 28 years). Five individuals possessed prior expertise with SSVEP-BCI, whereas the others were novices. Written consent was acquired [10]. The 12-class SSVEP dataset was acquired via a Biosemi Active Two EEG equipment equipped with eight occipital electrodes. Stimuli were encoded using frequencies ranging from 9.25 Hz to 14.75 Hz in increments of 0.5 Hz, and phases of 0 , 0.5π , π , and 1.5π . EEG data were initially collected at 2048 Hz and subsequently downsampled to 256 Hz. Each participant underwent 15 trials, each consisting of 12 randomly sorted targets shown in a 4×3 grid (6×6 cm) for 4 s, with a 135 ms visual delay.

V. METHODOLOGY

A. SPATIO-SPECTRAL CANONICAL CORRELATION ANALYSIS (SSCCA)

The spatio-spectral canonical correlation analysis (SSCCA) method, originating from the common spatio-spectral patterns (CSSP) technique, improves spatial and frequency information by delaying EEG input.

Canonical correlation analysis (CCA) maximizes the correlation between two sets of variables, Z and Y , utilizing weight vectors W and V . In this context, Z (of dimensions $N_c \times N_s$) and Y (of dimensions $N_c \times N_s$) are matrices comprising real values, with canonical variables defined as $H_z = W^T Z$ and $H_y = V^T Y$. The spatial filters W and V are engineered to synchronize the EEG and reference signals, hence optimizing their correlation ρ which may be computed using Eq. 1.

$$\rho = \arg \max_v \frac{\text{corr}(H_z, H_y)}{\sqrt{\text{var}(H_z)} \sqrt{\text{var}(H_y)}} \quad (1)$$

$$\rho = \arg \max_v \frac{W^T Z Y^T V}{\sqrt{W^T Z Z^T W} \sqrt{V^T Y Y^T V}} \quad (2)$$

The variables N_s and N_c denote the signal length and the quantity of EEG channels, respectively. Z represents the multi-channel EEG signal, while Y is the reference signal. Eq. 3 and 4 delineates the canonical variables comprehensively.

$$Z = \begin{bmatrix} z_1 \\ \vdots \\ z_{N_c} \end{bmatrix}, W = [w_1, w_2, \dots, w_{N_c}]^T, Y = \begin{bmatrix} y_1 \\ \vdots \\ y_{N_c} \end{bmatrix} \quad (3)$$

$$, V = [v_1, v_2, \dots, v_{N_c}]^T$$

$$H_z = w_1 z_1 + w_2 z_2 + \dots + w_{N_c} z_{N_c}, H_y = v_1 y_1 + v_2 y_2 + \dots + v_{N_c} y_{N_c} \quad (4)$$

To ascertain the SSVEP frequency using CCA, reference signals for each visual stimulus, represented as $[Y_{f^1}, Y_{f^2}, \dots, Y_{f^k}]$, are generated. In this context, k represents the quantity of stimuli, while f^k denotes the frequency of the k^{th} stimulus. CCA calculates correlation coefficients $[\rho_1, \rho_2, \dots, \rho_r]$, where r represents the minimum rank of the matrices Z (EEG signals) and Y (reference signals). The correlation results indicate a robust relationship between EEG and reference signals for each stimulus, especially between the SSVEP frequency and the reference. Time-delay embedding improves CCA efficacy by integrating both original and delayed EEG signals as inputs. Thus, Eq. 3 is modified in the following manner

$$\hat{Z} = \begin{bmatrix} Z \\ \delta_\tau Z \end{bmatrix}, \hat{H}_z = W_0^T Z + W_\tau^T Z(\delta_\tau Z) \quad (5)$$

The expanded form of the Eq. 5 and the canonical variable will be in the form of Eq. 6:

$$Z = \begin{bmatrix} z_1 \\ \vdots \\ z_{N_c} \\ \delta_\tau z_1 \\ \vdots \\ \delta_\tau z_{N_c} \end{bmatrix}, W = [w_1^0, w_2^0, \dots, w_p^0, w_1^\tau, w_2^\tau, \dots, w_{N_c}^\tau]^T,$$

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_{N_c} \end{bmatrix}, V = [v_1, v_2, \dots, v_{N_c}]^T \quad (6)$$

The canonical variables of the embedded EEG signal and the reference signal will be expressed in the form of the Eq. 7.

$$\hat{W}_z = W_0^T Z(n) + W_\tau^T Z(n - \tau) = v_1 y_1 + \dots + v_{N_c} y_{N_c} = V^T Y \quad (7)$$

Eq. 7 is similar to FIR filter, therefore, when calculating the canonical variables \hat{H}_x , The FIR frequency filter is utilized in conjunction with spatial filtering, whereas time delay embedding facilitates concurrent frequency and spatial optimization during the correlation computation. The correlation coefficient (ρ) shown in Eq. 1 and Eq. 2.

We may utilize the SSCCA to identify the frequency of the SSVEPs. Using the 2 formula N_c numbers of coefficients $\rho = [\rho_1, \rho_2, \dots, \rho_{N_c}]$. The maximum correlation coefficient is employed for frequency recognition, with the associated template signal frequency determined as that of the test signal.

$$f_{\text{test}} = \max_i(\rho_i) \quad (8)$$

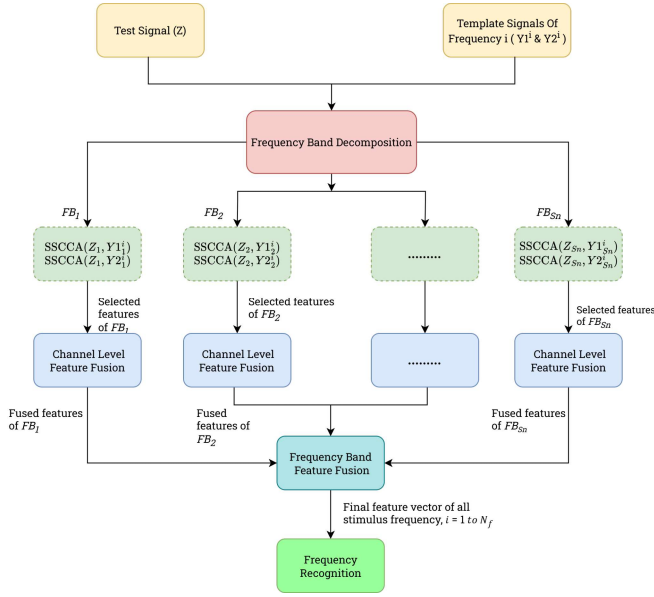


Fig. 1: Block Diagram of the Proposed Methodology

B. PROPOSED METHOD

This study presents a methodology based on SSCCA for the integration of spatio-spectral features to detect SSVEP frequencies. The framework comprises three stages: (1) Sub-band decomposition; (2) Feature fusion; and (3) Frequency recognition are elaborated upon in the subsequent subsections.

1) **SUBBAND DECOMPOSITION:** Initially, bandpass filters are employed to partition data into discrete subbands. The dataset comprises four elements: 8 channels, 1041 sample points, 15 trials, and 12 stimuli, resulting in a $N_c \times N_s \times N_t \times N_f$ array. The stimulus frequency varies from 9.25 Hz to 14.75 Hz, with an interval of 0.5 Hz. The utilization of filterbank technology enhances BCI performance. Each trigger undergoes processing using a Chebyshev Type I filter (12th order, 3dB stopband attenuation) to generate S_n subbands, each characterized by unique harmonic frequencies and a shared cutoff of 80 Hz. The lower cutoff frequency of the m^{th} bandpass filter is denoted as $m \times f_0$, where f_0 signifies the initial stimulus frequency and $m = 1, 2, \dots, S_n (S_n = 5)$.

Z denotes the test signal, whereas $Y1^i$ and $Y2^i$ signify the template signals at the i -th frequency in $\mathbb{R}^{N_c \times N_s}$. $Y1^i$ represents the mean of the first segment (trial 1 to trial 7) of the training trials, while $Y2^i$ denotes the mean of the latter segment (trial 8 to trial 14). Subsequent to filtering, the test signal is partitioned into frequency subbands Z_1, Z_2, \dots, Z_{S_n} . Correspondingly, the template signal is partitioned into subbands $Y1_1^i, Y1_2^i, \dots, Y1_{S_n}^i$ and $Y2_1^i, Y2_2^i, \dots, Y2_{S_n}^i$.

2) **FEATURE FUSION:** During the frequency band feature fusion, features are recovered using SSCCA to assess the correlation between each pair of subband signals from the test and template signals. The quantity of features is equivalent to the quantity of channels, N_c .

For each subband, two coefficient vectors are computed utilizing SSCCA: one derived from the test signal Z_m and the

first template signal $Y1_m^i$, and the other from Z_m and the second template signal $Y2_m^i$. The index m denotes the subbands ($m = 1, 2, \dots, S_n$). Each pair produces N_c coefficients, which are consolidated into $2 \times N_c$ coefficients. Upon arranging them in descending order, the first N_c coefficients are chosen as the feature vector. In the initial phase of feature fusion, individual coefficients from SSCCA are amalgamated, with the template signal established as the average of all training trials (trial 1 to trial 14).

Fig. 2a illustrates the correlation coefficient index in relation to average classification accuracy among individuals for four temporal intervals (0.25s, 0.50s, 0.75s, and 1.00s), exhibiting a non-linear decline in accuracy.

The non-linear weighting function ϕ depicted in Fig. 2b improves system performance through the integration of coefficients.

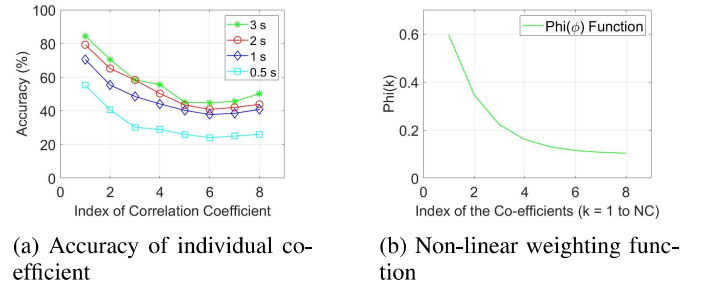


Fig. 2: Accuracy of individual coefficient and Non-linear weighting function in channel level feature fusion.

This study utilizes a channel level feature fusion methodology, amalgamating coefficients $\phi_1, \phi_2, \dots, \phi_{N_c}$ through a non-linear weighting function $\phi_k = e^{-a_1 k} + b_1$, where $k = 1, 2, \dots, N_c$, $a_1 = 0.6$, and $b_1 = 0.2$. Grid search optimizes these parameters. The correlation coefficients, obtained using SSCCA, signify distinct occipital brain regions and are amalgamated to create the m^{th} subband feature. Each subband coefficient encompasses essential information for identifying stimulus frequencies. Fig. 3a depicts the correlation between frequency band indices and mean classification accuracy throughout four temporal periods (0.25s, 0.50s, 0.75s, and 1.00s), demonstrating a persistent non-linear decrease in accuracy.

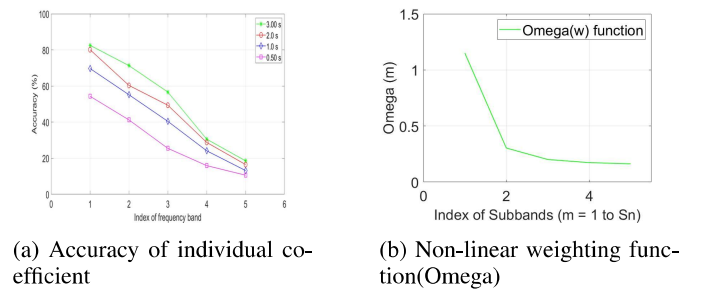


Fig. 3: Accuracy of individual coefficient and Non-linear weighting function in frequency band feature fusion.

Fig. 3b illustrates the non-linear weighting function (Ω) employed in the frequency band feature fusion to improve system performance through the amalgamation of coefficients.

The combined feature for the test signal Z is produced using the template signals Y^1_i and Y^2_i at the i -th frequency ($i = 1, 2, \dots, N_f$) and is calculated as:

$$\psi_i = \sum_{m=1}^S w_m \cdot \delta_m^i \quad (9)$$

The non-linear weighting function is defined as $w_m = m^{-a_2} + b_2$, with $m = 1, 2, \dots, S_n$, $a_2 = 2.0$, and $b_2 = 0.25$. The settings are tuned by grid search. This process generates fused features ψ_i for each frequency i .

3) **FREQUENCY RECOGNITION**: The concluding process involves determining the stimulus frequency. The attributes acquired from the preceding channel level and frequency band feature fusion are utilized for each stimulus frequency, $\psi_1 \dots \psi_2 \dots \psi_{N_f}$. The equation calculates the frequency F_z of the test signal Z .

$$F_z = \max_f \psi_i(f), i = 1, 2, \dots, N_f \quad (10)$$

VI. RESULT & PERFORMANCE ANALYSIS

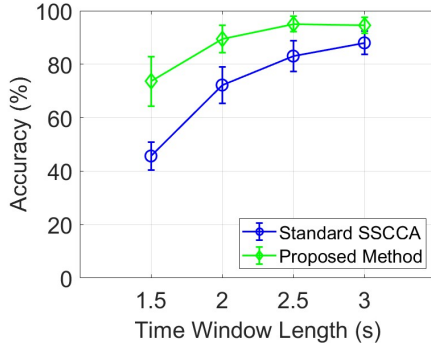


Fig. 4: The average frequency recognition accuracy of SSCCA and the suggested technique among individuals at different time intervals.

The efficacy of the proposed method was assessed using precision as the evaluation parameter. The average accuracy during a 3s interval was compared to the traditional SSCCA approach, as illustrated in Fig. 4. The findings, accompanied by error bars representing standard errors, illustrate that the suggested SSCCA spatial filtering regularly surpasses the conventional SSCCA, with a peak accuracy of 94.5% with a 3s time window and a 1s delay.

The suggested model attains an impressive accuracy of 94.5% (Table I), illustrating its reliability and efficacy in SSVEP frequency recognition, as well as its potential to improve the performance of real BCI systems.

VII. CONCLUSION

This study introduces a multistage feature fusion framework utilizing SSCCA-based spatial filtering for precise SSVEP

TABLE I: Performance Comparison of Different Models

Model	Accuracy
Standard CCA [10]	55%
MwayCCA [10]	68.39%
LI-MCCA [10]	70.28%
MsetCCA [11]	73.61%
UD-C-CNN [12]	92.33%
CORRCA [11]	81.44%
HSCA [13]	88%
Proposed Method	94.5%

frequency detection in brain-computer interfaces. Although it enhances accuracy and dependability, its disadvantages encompass dependence on training data, manual analysis, constrained stimuli, and reliance on grid search for parameter optimization. Future endeavors will tackle these issues by augmenting datasets, incorporating recordings, diversifying stimulation frequencies, and creating cost-effective, real-time systems to improve the accessibility and practicality of SSVEP-based BCIs.

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] 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*, vol. 53, no. 12, pp. 2610–2614, 2006.
- [3] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in ssvep-based bci using multiset canonical correlation analysis," *International journal of neural systems*, vol. 24, no. 04, p. 1450013, 2014.
- [4] Y. Jiao, Y. Zhang, Y. Wang, B. Wang, J. Jin, and X. Wang, "A novel multilayer correlation maximization model for improving cca-based frequency recognition in ssvep brain-computer interface," *International journal of neural systems*, vol. 28, no. 04, p. 1750039, 2018.
- [5] M. N. Cherloo, H. K. Amiri, and M. R. Daliri, "Spatio-spectral cca (ss-cca): a novel approach for frequency recognition in ssvep-based bci," *Journal of Neuroscience Methods*, vol. 371, p. 109499, 2022.
- [6] K. Qin, R. Wang, and Y. Zhang, "Filter bank-driven multivariate synchronization index for training-free ssvep bci," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 934–943, 2021.
- [7] Y. Güçlütürk, R. H. Jacobs, and R. v. Lier, "Liking versus complexity: Decomposing the inverted u-curve," *Frontiers in human neuroscience*, vol. 10, p. 173211, 2016.
- [8] S. Mahmood, J. Shin, I. Farhana, M. R. Islam, and M. K. I. Molla, "Frequency recognition of short-time ssvep signal using corcca-based spatio-spectral feature fusion framework," *IEEE Access*, vol. 9, pp. 167 744–167 755, 2021.
- [9] V. Mishuhina and X. Jiang, "Feature weighting and regularization of common spatial patterns in eeg-based motor imagery bci," *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 783–787, 2018.
- [10] M. Nakanishi, Y. Wang, Y.-T. Wang, and T.-P. Jung, "A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials," *PloS one*, vol. 10, no. 10, p. e0140703, 2015.
- [11] N. Waytowich, V. J. Lawhern, J. O. Garcia, J. Cummings, J. Faller, P. Sajda, and J. M. Vettel, "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *Journal of neural engineering*, vol. 15, no. 6, p. 066031, 2018.
- [12] A. Ravi, N. H. Beni, J. Manuel, and N. Jiang, "Comparing user-dependent and user-independent training of cnn for ssvep bci," *Journal of neural engineering*, vol. 17, no. 2, p. 026028, 2020.
- [13] R. Miao, L. Zhang, and Q. Sun, "Hybrid template canonical correlation analysis method for enhancing ssvep recognition under data-limited condition," in *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2021, pp. 65–68.