# Multiclass Steady-State Visual Evoked Potential Frequency Evaluation using Chirp-Modulated Stimuli

Nicholas R. Waytowich, *Member, IEEE,* Dean J. Krusienski, *Senior Member, IEEE*

***Abstract*—Steady-State Visual Evoked Potentials (SSVEPs) are oscillations of the electroencephalogram (EEG) observed over the occipital area that exhibit a frequency corresponding to a repetitively flashing visual stimulus. SSVEPs have proven to be very consistent and reliable signals for rapid EEG-based brain-computer interface (BCI) control. While a subject-specific SSVEP stimulus frequency optimization is ideal, this can be a tedious and time-consuming process. Thus, many studies select SSVEP stimulation frequencies somewhat arbitrarily. Not only is there no standardized set of SSVEP stimulus frequencies or frequency selection method, but some studies even claim conflicting frequency ranges for optimal performance. In this work, 17 subjects were stimulated with an LED array that flashed according to a chirp-modulated signal having a frequency that varied linearly over the typical functional range of SSVEP. The resulting EEG was analyzed using canonical correlation analysis (CCA) and a genetic algorithm (GA) was implemented to determine generalized stimulation frequency sets over a continuum of simulated multiclass BCI classification scenarios. The results show that distinct frequency feature groupings exist over the different multiclass scenarios, and that these groupings result in different information transfer rates. These offline results can provide a guide for generalized stimulus frequency selection for SSVEP-based BCIs with an arbitrary number of targets.**

***Index Terms*—Steady-State Visual Evoked Potentials, Brain-Computer Interfaces, Canonical Correlation Analysis, Genetic Algorithms, Feature Selection.**

## I. Introduction

BRAIN-COMPUTER INTERFACES (BCIs) are assistive devices that analyze brain activity and decode user intent in order to provide a non-neuromuscular pathway of communication and control [1]. Some of the most promising approaches for scalp-recorded EEG-based BCIs utilize Steady-State Visual Evoked Potentials (SSVEPs), which are oscillations in EEG that correspond to the frequency of a flashing visual stimulus [2][3]. The SSVEP response can be elicited with flashing stimuli in the range of 1-100 Hz [4]. However, most SSVEP BCI studies utilize flashing stimuli from the 5-45 Hz range [5], [6], [7], [8].

Due to physiological differences between subjects, there tends to be high intersubject variability in terms of SSVEP response power at different stimulation frequencies. While a subject-specific SSVEP stimulus frequency optimization is ideal, this can be a tedious and time-consuming process. Thus, many studies select SSVEP stimulation frequencies somewhat arbitrarily. Not only is there no standardized set of SSVEP stimulus frequencies or a standardized frequency selection method, but some studies even claim conflicting frequency ranges for optimal performance. A detailed review of SSVEP stimulation techniques by Zhu et al. [8] showed that most studies not only varied greatly in the stimulus frequencies used, but also varied in the number of classes used. The majority of studies did not perform stimulus frequency optimization prior to the study, nor provided a justification for the selected frequencies. This generally arbitrary selection of stimulus frequencies leads to suboptimal fixed-frequency results (i.e., not subject-specific) and makes performance comparisons across SSVEP studies difficult. Several recent studies have performed fairly comprehensive evaluations of various SSVEP stimulus parameters, including stimulus size, shape, color, and pairwise comparisons of stimulus frequencies based on signal-to-noise ratio [9], [10]. However, these studies did not examine optimal sets of frequencies for multiclass SSVEP applications.

Recently, multichannel SSVEP-BCIs have been demonstrated online using Canonical Correlation Analysis (CCA), which can achieve high information transfer rates with little to no training data [11], [12]. CCA is generally a preferred detection method for SSVEP BCIs because of its inherent channel selection and harmonic analysis capabilities, relative simplicity, and robust performance. Whereas an earlier study explored individualized optimization of SSVEP stimulus frequencies using CCA [13], this study uses the same data set to investigate generalized stimulation frequency sets over a continuum of simulated multiclass BCI classification scenarios. Subjects were visually stimulated with an LED array that flashed according to a chirp-modulated signal [14] having a frequency that varied linearly over the typical functional range of SSVEP from 5.5-34.5 Hz. The resulting EEG was analyzed using canonical correlation analysis (CCA) and a genetic algorithm (GA) was implemented to determine the generalized stimulation frequency set that maximizes classification performance for numbers of classes (i.e., simulated flashing target groupings) ranging from 2 to 36.

## II. Methodology

### A. Data Collection

EEG data were collected from 22 healthy volunteers (5 women, 17 men; age range 18-42 years) from a single session. All subjects were free of any neurological or psychiatric disorders and had either normal or corrected-to-normal vision. Each subject gave written informed consent prior to participation

N.R. Waytowich and D.J. Krusienski are with the Biomedical Engineering Program, Old Dominion University, Norfolk, VA, 30332

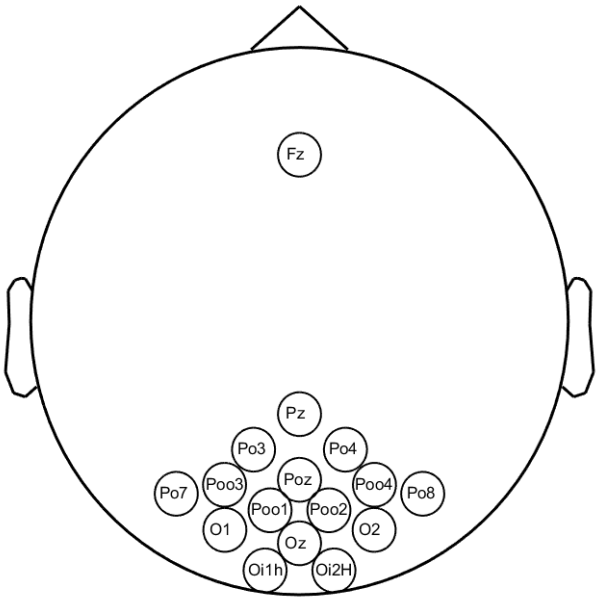



Fig. 1. The EEG electrode montage used for data collection. The positions are based on the International 10-20 system.

and all aspects of the study were reviewed and approved by the Old Dominion University Institutional Review Board.

After a screening of the EEG data for discriminable SSVEP responses using a visual evaluation of the subject-specific CCA characterization similar to Figure 2, five subjects were excluded because they produced little or no detectable SSVEP response. Thus, 17 (3 women, 14 men; age range 18-36 years) of the 22 subjects were included in the subsequent analysis.

EEG data were recorded using 16 active electrodes and a g.USBAmp biosignal amplifier (g.tec Medical Engineering). Electrodes were positioned primarily over the occipital and parietal regions at locations based on the International 10-20 system [15]: Fz, Pz, POz Oz, O1, O2, PO3, PO4, PO7, PO8, POO1, POO2, POO3, POO4, OI1h, OI2h, as shown in Figure 1. All EEG data were bandpass filtered from 0.1 Hz to 100 Hz, notch filtered at 60 Hz, and digitized at 512 Hz. Data recording and timing were controlled by BCI2000 general-purpose BCI software [16].

### B. Experimental Paradigm

Each subject sat in a dark room in front of a custom-built SSVEP stimulator composed of an 8 x 8 array of green LEDs with dimensions of 5.84 cm x 5.84 cm. Each LED in the array was wired together so that all LEDs illuminated simultaneously with the preprogrammed stimulus. The LED stimulator was placed in the center of the subjects' visual field approximately 60 cm away so that the stimulation spanned visual angles of 5.25 degrees vertically and horizontally. The stimulator was driven by an Arduino Mega microcontroller board with an output stimulation frequency of 500 Hz and a 10-bit intensity resolution. LED luminosity was linearized over the operating range to ensure a uniform intensity distribution, and the LED array was tested using a photo-diode to verify consistent stimulation frequencies. All stimulation signals were generated using Matlab software (Mathworks, Nattick, MA) and loaded onto the microcontroller before stimulation.

During the experiment, the subject's task was to visually attend and keep focus on the flashing stimulator. During the course of a single session, subjects performed several runs of attending to 30 seconds of continuous stimulation followed by 15 seconds of rest. Runs were repeated five times for each of four different chirp-modulated stimulus waveforms. The waveforms that were presented were composed as follows: 1) square waveform with a chirp increase from 5.5-20.5 Hz, 2) square waveforms with a chirp increase from 19.5-34.5 Hz, 3) square waveform with a chirp decrease from 20.5-5.5 Hz, and 4) square waveform with a chirp decrease from 34.5-19.5 Hz. Each chirp-modulated waveform (increase and decrease) had a chirp rate of $\Delta_f$ = 0.5 Hz per second. This provided approximately two seconds centered on each integer frequency and is sufficiently slow enough to emulate a fixed frequency over a short time window.

A total of 5 trials per waveform were presented, giving a total of 20 trials. The waveforms were presented in a counterbalanced order to account for possible fatigue issues. Baseline EEG with the stimulator turned off and the subjects' eyes kept open was collected for approximately two minutes before and after the session. Each subject participated in a single experimental session lasting approximately 30 minutes.

### C. Canonical Correlation Analysis

Canonical Correlation Analysis (CCA) is a multi-dimensional statistical analysis technique that finds underlying linear correlations between two sets of data. Given two multi-dimensional data sets $X$, and $Y$, linear combinations $x = X^T W_x$ and $y = Y^T W_y$ can be found that maximize the correlation between $x$ and $y$. The CCA finds the weight vectors $W_x$ and $W_y$ by solving the following optimization problem:

$$\max_{W_x, W_y} \frac{E[W_x^T X Y^T W_y]}{\sqrt{E[W_x^T X X^T W_x] E[W_y^T Y Y^T W_y]}} \tag{1}$$

In practice this can be solved using the singular-value decomposition method to diagonalize the covariance matrices as the maximum canonical correlation corresponds to the square-root of the largest eigenvalue.

Lin et al. [11] used CCA for SSVEP classification. In this case, CCA is used to find linear correlations between multichannel EEG data, $X$, and a set of reference signals $Y_f$. A reference set is made for each SSVEP stimulus that consists of sine and cosine signals modulated at the fundamental and harmonic frequencies of the stimulus. The reference signal $Y_f$, shown below, can be derived using $N_h$ harmonics.

$$Y_f = \begin{pmatrix} \sin(2\pi f t) \\ \cos(2\pi f t) \\ \vdots \\ \sin(2\pi N_h f t) \\ \cos(2\pi N_h f t) \end{pmatrix} \tag{2}$$

EEG data is canonically correlated with each reference signal and the classification output is determined as $f_s = \max_i \rho(f)$,

where $f = f_1, f_2, ...f_K$ and $K$ is the total number of classes (target frequencies) in the SSVEP BCI.

### D. Data Preprocessing

All SSVEP data was first bandpass filtered using a zero-phase, IIR filter from 0.5-40 Hz. All data were segmented by trial for each chirp stimulus condition and inter-trial data (i.e. rest periods) were discarded. The data elicited from the chirp signals were then analyzed using the CCA method. Reference signals $Y_f$ were created and centered around each frequency from 6 to 33.5 Hz spaced every 0.5 Hz for a total of 56 reference signal sets that spanned the chirp signal. All reference signals were created using only the fundamental frequency $(N_h = 1)$ for this study.

The chirp data from the 5.5-20.5 Hz waveform was concatenated with the chirp data from the 19.5-34.5 Hz waveform to produce an SSVEP response signal from 5.5-34.5 Hz. The decreasing chirp signals (20.5-5.5 Hz and 34.5-19.5 Hz) were concatenated to produce a 34.5-5.5 Hz response and then time-reversed to match the previous 5.5-34.5 Hz. Fixed frequencies were approximated from the chirp signal by using a sliding window with a length of two seconds and a one-second overlap. The data representing the 34-34.5 Hz was discarded due to a recording inconsistency at the beginning or end of the decreasing or increasing chirp trials, respectively. Thus, the frequency features used in this study range from center frequencies of 6 Hz to 33.5 Hz in increments of 0.5 Hz (56 total distinct frequencies). This covers the functional range of the SSVEP spectrum with sufficient resolution.

For all 17 subjects, the CCA was performed on each window of the EEG response and on each frequency from the SSVEP reference signals. Each target frequency (i.e. the current time window corresponding to the frequency from the chirp stimuli) was canonically correlated with each reference signal frequency. This results in a quantitative measure of target discrimination from background EEG activity for each of the selected frequencies (shown in Figure 2).

The data from the 17 subjects were combined into a single set consisting of 170 observations per frequency. The combined subject set is used to build an average SSVEP classifier for the optimization to produce a generalized solution for each N-class condition.

### E. Feature Selection using a Genetic Algorithm

Genetic algorithms (GAs) are powerful search algorithms that simulate natural evolution for optimization problems [17]. GAs have been used for wrapper-based feature selection problems and have been utilized to optimize feature selection in BCIs [18]. In this study, the GA is used to find optimal sets of SSVEP frequency features using CCA for a range of simulated N-class classification problems. The N-class conditions simulated in this study range from 2 to 36 classes in order to compare the performance between practical class sizes and determine generalized feature sets that optimize performance for each N-class condition. Each 0.5 Hz increment of the chirp stimulus ranging from 6 Hz to 33.5 Hz was used to generate a simulated frequency feature. This resulted in 56

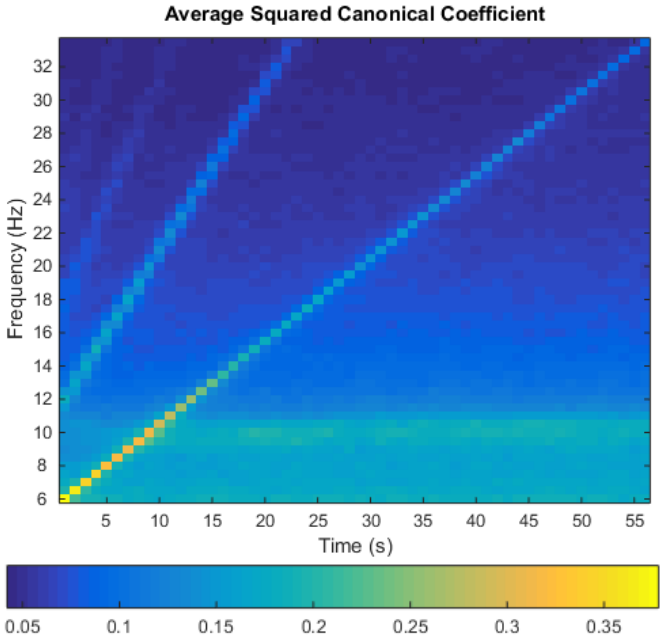



Fig. 2. Canonical Correlation Analysis (CCA) squared correlation values averaged across the 17 subjects. The x-axis shows the timescale based on the chirp stimulus. Each column represents a 2-second segment of the chirp signal that is approximated as a fixed frequency. The strong diagonal correlations represent a correspondence of the EEG and the stimulation frequency and its harmonics across the windowed time segments.

total frequency features, which constitutes an extremely vast search space for feature selection that is not practical for exhaustive evaluation for most of the N-class conditions.

*1) Initial Population:* An independent GA optimization was performed for each of the 35 N-class BCI conditions (i.e., 2-36 classes). For each condition, an initial population of 1000 individuals (candidate solutions) is created. Each individual is a binary string of length 56, where the ones represent the features that are to be included in the model and zeros represent the features to disregard. For each N-class condition, the number of selected features is constrained to match N, as each feature represents a frequency target.

For the initial population, the majority of individuals are randomized with N ones and (56-N) zeros using a uniform random distribution. After 50 generations, the population was reinitialized with uniform random distribution and 5-10% of the population was seeded with a combination of the 10 best solutions from the previous GA optimization and random 1-bit permutations of these solutions. These permutations were selected to minimize the Hamming distance from the best solutions since the best solutions were generally found to be completely or partly contiguous over frequency spans. This was done in order to fine tune the search near the best solution while still maintaining diversity of the search.

*2) Fitness Evaluation:* The fitness of each individual is evaluated by classifying the EEG using the CCA algorithm. Each individual is assigned a fitness value based on its respective classification performance. In this study, two fitness evaluation schemes are compared: the overall classification error and empirical confusion matrix norm (ECMN).

**Classification Error Optimization:** The GA was tuned to select and reproduce individuals with features that minimize the classification error from the CCA classifier. The overall classification error metric is calculated from the confusion matrix as:

$$\text{classification error} = 1 - \frac{\Sigma_i H_{ii}}{N_C} \tag{3}$$

where $H_{ii}$ is the number of correctly classified observations for class $i$ and $N_C$ is the total number of observations in the confusion matrix.

**Empirical Confusion Matrix Norm Optimization:** This approach minimizes the norm of the empirical confusion matrix (ECM) $||C_s||$ [19], [20]. For this scheme, the diagonals (correct classifications) of the standard confusion matrix are zeroed out to create the empirical confusion matrix $C_S$. The ECMN is calculated as:

$$||C_S|| = \sqrt{\lambda_{max}(C_S^* C_S)} \leq \sqrt{Tr(C_S^* C_S)} \tag{4}$$

where $C_S^*$ denotes the conjugate transpose of $C_S$, $\lambda_{max}$ is the maximum eigenvalue and $Tr()$ is the sum of the diagonal elements. The latter equation constitutes the upper bound of the confusion matrix norm. The value of $||C_S||$ will be higher for confusion matrices with miss-classifications that bias towards a certain set of classifier outputs. Unlike classification error optimization, which seeks to find the set of SSVEP features that provide the overall best classifier performance, the confusion norm optimization seeks features that provide the best classifier balance. This is particularly important for practical BCI implementations since lower frequencies tend to produce responses with higher signal-to-noise ratios (SNRs) that can dominate the classifier output, which may be masked for higher class conditions if only the overall accuracy is evaluated.

*3) Selection and Reproduction:* For the GA selection procedure, the population is ranked and sorted according to the individuals' fitness, and the top 150 individuals are labeled as the elite group. Individuals from the elite group are randomly selected for mating with a randomly selected member of the entire population (i.e., at least one parent is always from the elite group). This elitist procedure is done to speed convergence of the GA [21]. Thus, 850 children are generated in each generation of the GA.

*4) Crossover and Mutation:* The crossover of two parent genes (i.e., binary strings) to form a child gene is performed using bitwise logical AND and bitwise logical OR. Given the binary strings from two parents, the bitwise OR operation is used to select the feature from each parent providing a total of $R$ features that will be transferred to the child gene. This results in $R \geq N$ and thus an $n$ subset of $R$ must be selected. The bitwise AND is then used to select the $S$ features that are common to both parents. If $S < N$ then the remaining $N - S$ features are randomly selected from the original set of $R$ features. The final child has $N$ features derived from the parent features. A mutation rate of 5% was applied to the children genes after reproduction and crossover. For each generation, 5% of the children's genes were randomly selected for mutation where one random bit was flipped from 1 to 0 and another random bit was flipped from 0 to 1 to maintain the desired number of features.

*5) Batch Processing and Evaluation:* The combined EEG data set from the 17 subjects was used for the GA optimizations. A separate GA optimization was performed for each of the 35 different N-class conditions (2 through 36-class) for both the classification error and ECMN optimization, resulting in a grand total of 70 distinct GA optimizations. The runtime of each GA was fixed to 100 generations (i.e., 50 generations for determining the best seeds and 50 generations for fine tuning the solution) to provide a sufficient search. For each GA evaluation, the features and fitness values were stored for each individual and generation to evaluate the convergence and sensitivity of the solutions.

The resulting performances are quantified using the overall classification accuracy and the corresponding information transfer rate (ITR). The ITR, or bitrate, is a common metric for evaluating performance of a BCI that accounts for speed, accuracy, and class-size. The ITR in bits/min is computed as:

$$ITR = \left( \log_2 N + P \log_2 P + (1-P) \log_2 \left( \frac{1-P}{N-1} \right) \right) * \frac{60}{T} \tag{5}$$

where $N$ is the number of possible classes, $P$ is the classification accuracy and $T$ is the selection time used to obtain the classification. The nominal selection time is 2 seconds based on the length of each SSVEP observation used in this study. In order to simulate a practical ITR, an inter-epoch duration of 0.3 seconds was utilized based on the recent SSVEP study from [5] which estimates the time required to shift eye-gaze from one target to the next. These simulated ITRs are intended to provide a relative performance comparison between the optimization methods rather than a completely realistic estimate of ITR.

## III. Results

To compare the GA solutions for each of the N-class conditions, the performance results are plotted as line plots for each N-class condition. The GA solutions (best feature set) for both the classification error and the ECMN optimization approaches are compared in Figure 3. The resulting overall accuracy and ITR of the best individual from both approaches are plotted against each of the N-class conditions. The results in Figure 3 show the expected decrease in accuracy of the optimization solutions for increased number of classes. All accuracies produced are significantly higher than chance. The accuracy of the best solution from the ECMN optimization shows a slight decrease in accuracy compared to the best solution from the classification error optimization. The drop in accuracy only becomes apparent after a class size of 9 and exists until an class size of 35. The ITR plots in Figure 3 show an even more pronounced degradation in performance of the ECMN optimization compared to the classification error optimization for class sizes 9-34. For both optimization conditions, an inflection point in accuracy and ITR is observed around 9-10 classes, indicating suboptimal performance for these conditions. In order to quantify and compare the accuracy imbalance across target frequencies, the ECMN for both

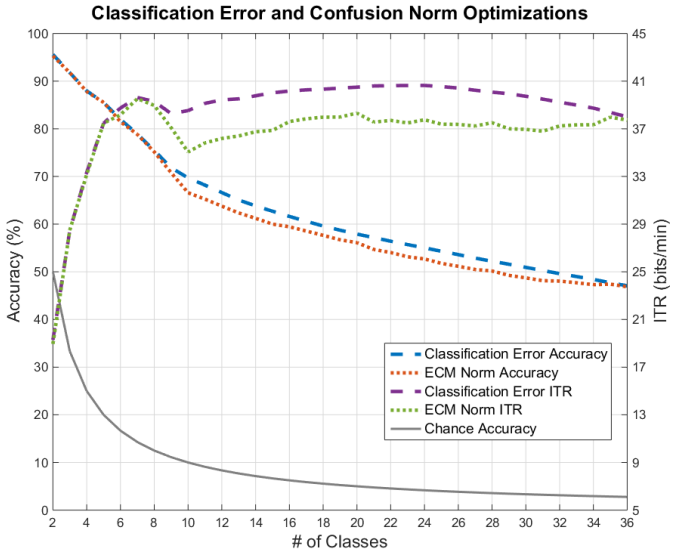



Fig. 3. Accuracy and Information Transfer Rate (ITR) of the best solution from the classification error and ECMN optimizations for each N-class condition. The red and blue accuracy line plots correspond to the left y-axis scale and the green and purple ITR line plots correspond to the right y-axis scale. Chance accuracy for each N-class condition is plotted as the solid gray line.

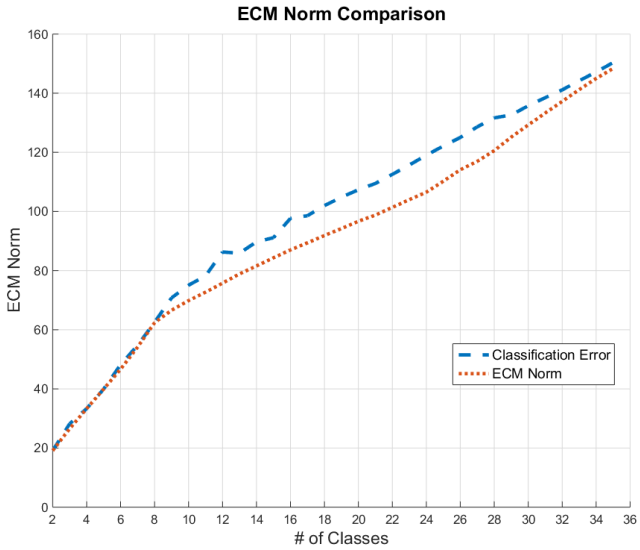
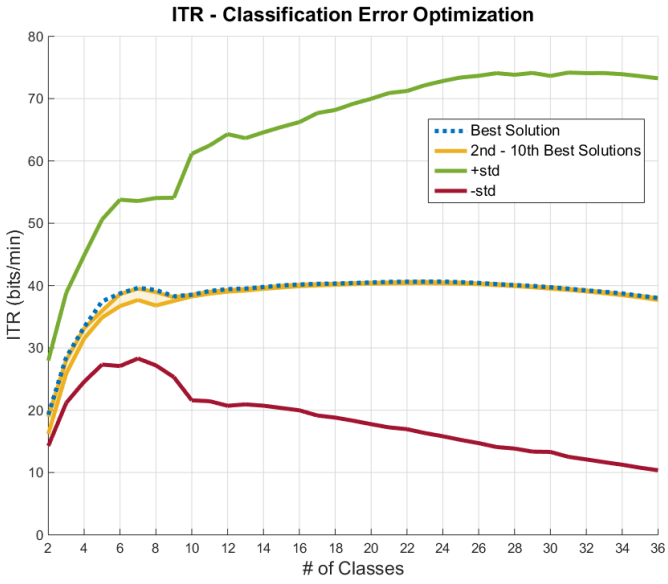
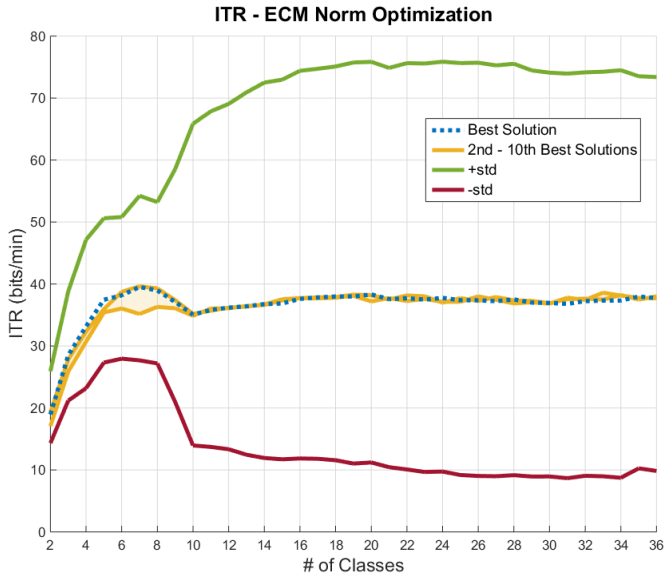



Fig. 4. A comparison of the ECM Norm for the best solutions selected by the Classification Error and ECM Norm optimizations, respectively. The ECM Norm provides a measure of classification accuracy imbalance across target frequencies, where a lower ECM Norm indicates more balanced accuracies across target frequencies.




(a)




(b)

Fig. 5. Information Transfer Rate (ITR) performance for the top 10 best optimization solutions for each N-class condition. (a) shows the top 10 results from classification error optimization and (b) shows the top 10 results from the ECMN optimization. For each plot, the dotted line indicates the best solution and the solid lines bound the 2nd-10th best solutions. The +/-std indicates the standard deviation of accuracy across subjects converted to ITR.

approaches is shown in Figure 4, where a lower ECM Norm indicates more balanced accuracies across target frequencies. Similar to accuracy and ITR in Figure 3, it is observed that the differences in approaches occur for class sizes 9-34.

To illustrate the level of solution uniqueness and sensitivity from each GA evaluation, the ITR of the solution is plotted with the ITRs of the 10 next-best solutions in the population history as shown in Figure 5. Figure 5(a) shows the top 10 solutions for each N-class condition from the classification error optimization and Figure 5(b) shows the top 10 solutions for each N-class condition from the ECMN optimization. The performance difference between the best solution and the 9 next-best solutions is most apparent for the lower N-class conditions between 2 and 10. The higher N-class conditions show very similar ITR performances between the best and 9 next-best solutions. For the ECMN optimization in Figure 5(b), the next-best solutions sometime show a slight increase in ITR for a few of the higher N-class conditions. This is due to the fact that the ECMN is not optimizing overall accuracy, but the norm of the empirical confusion matrix.

Figure 6 shows the selected features from GA optimiza-

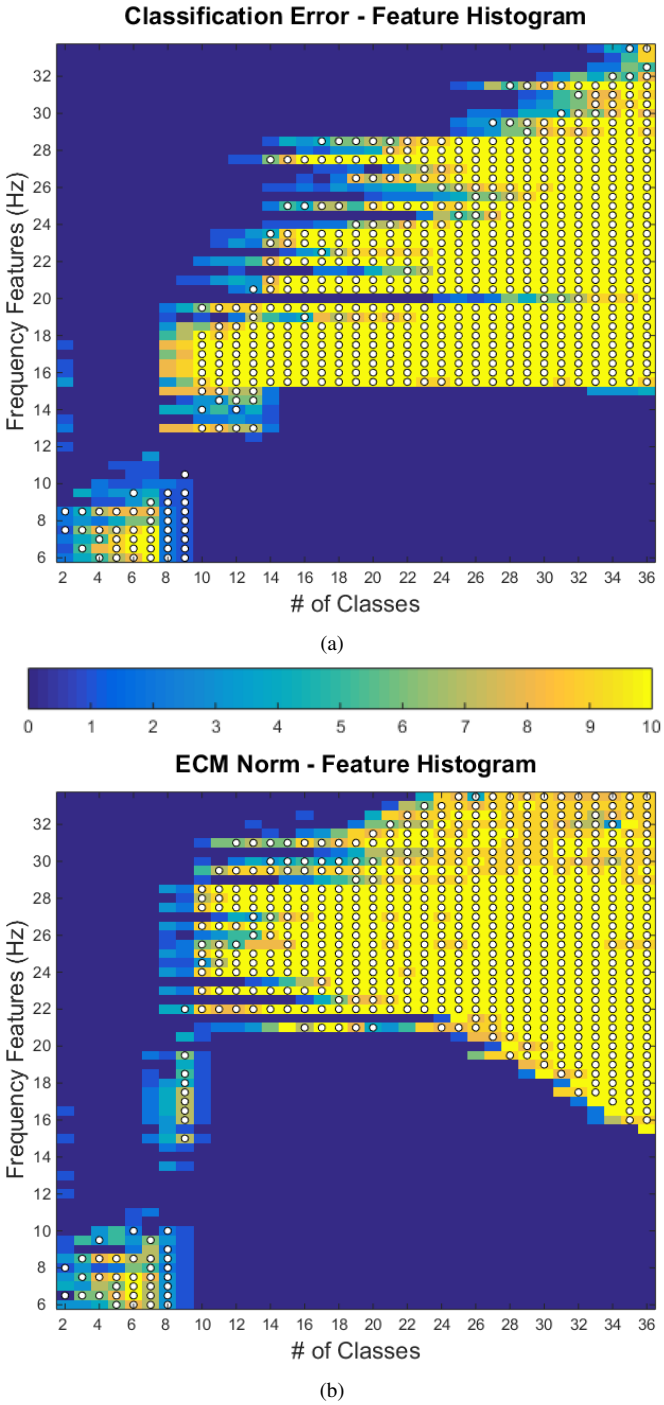



Fig. 6. 2-D histograms of optimized feature sets for each N-class condition. Each column is a different N-class condition that spans the possible feature range (6-33.5 Hz). For a given N-class condition, the features from the best optimized solution are superimposed as white circles indicating the selected feature. Each column has N dots corresponding to each N-class condition. The 9 next-best solutions are plotted as histograms where the color corresponds to the occurrence rate of the selected feature. (a) shows the feature distributions for the classification error optimization and (b) shows the feature distributions for the ECMN optimization.

tion for each N-class condition as 2-dimensional histograms. Figure 6(a) shows the selected feature distribution from the classification error optimization and Figure 6(b) shows the selected feature distribution from the ECMN optimization. Each column represents an N-class condition (i.e., number of simulated SSVEP targets) and each row represents a possible frequency feature. For each GA evaluation, the best set of features is plotted using white dots where the number of dots corresponds to the N-class condition and the location of each dot corresponds to the selected frequency feature. A histogram distribution of the selected features from the 9 next-best solutions for each N-class condition is plotted with a color scale corresponding to the number of times each frequency feature was selected for the next-best solutions. This gives a rough indication of the usefulness of a given feature with respect to the N-class condition.

## IV. Discussion

The feature selection histograms for both optimization methods in Figures 6(a) and 6(b), respectively, indicate the optimal features for lower N-class conditions come from lower frequencies. In both figures, there is a distinct change in selected features from the 6-9 Hz range to the 15-20 Hz range as the N-class condition increases from 8 to 9, which is also near the performance inflection point. This can be explained by examining the CCA plot in Figure 2. This plot shows that the frequencies from 6-10 Hz have a higher CCA coefficient across this entire band compared to the diagonals at higher frequencies. This is partly due to the higher power and SNR in this low frequency range, particularly in the alpha band. It should be noted that the alpha band and its second harmonic are generally avoided in the frequency selection due to the innate nature of this rhythm, independent of stimulation in this range.

Of further interest, no N-class conditions that are greater than N=9 contain any low frequency features in the final optimized set. For the classification error optimization, Figure 6(a) shows a bias towards the middle frequency range (15-19.5 Hz) which gradually includes higher frequencies as the N-class condition increases. However, for the ECMN optimization, Figure 6(b) shows another sharp change in selected features from the mid range (15-19.5 Hz) to the higher range (22-32 Hz) at the 10-class condition. This is likely due to the fact that these ranges of frequencies have more similar SNRs. This would result in a more balanced classification (i.e., minimizing the ECM norm), which is desirable for practical BCIs.

The contiguous nature of the majority of the selected feature sets can also be, in part, explained by the relative SNR across particular frequency bands. For the classification error, it is observed that the features are generally contiguous and after the inflection point, the added features come from the next highest frequency (i.e., next highest SNR). The ECMN optimization follows a similar contiguous pattern, except the initial range after the inflection point is (22-28 Hz) and roughly alternating lower and higher frequency contiguous features are included to maintain the class balance. Another contributing factor to the observed contiguous ranges is that

the harmonic frequencies play a role in the CCA as depicted in Figure 2. Particular frequency features, especially in the lower frequencies, will tend to dominate the CCA at the respective harmonic frequencies. Thus, it is unlikely for frequencies to be selected that have competing harmonics. Future work will examine the inclusion of harmonic frequencies in the CCA and its impact of on the selected feature sets.

In terms of maximizing the ITR across the population, the results indicate that a peak ITR near 40 bits/minute is achieved for the seven-class condition for both optimization schemes. Additionally, both optimization schemes selected contiguous frequency features from 6-9 Hz (0.5 Hz increment) for this condition. Aside from the inflection point, the ITR is generally flat above the 10-class condition due to the steady decrease in accuracy for higher classes. It should be noted that a fixed inter-stimulus interval was used across N-class conditions to compute the simulated ITRs. It is likely that longer inter-stimulus intervals would be needed for adequate scanning time in the higher-class conditions, which would further degrade the ITR as the number of classes increases. Figure 4 shows the expected improvement in balanced classification across target frequencies for the ECMN optimization for class sizes 9-34. Ultimately, if all target frequencies should be equi-probable for a given interface design, the ECMN optimization will provide a more balanced scheme for more than 8 classes with marginal decreases in overall classification performance.

The proposed approach has some limitations in terms of generalizing to an online BCI. The SSVEP responses were collected serially using a fairly large single-stimulus LED array. For an actual multi-class online implementation, simultaneous stimuli at different frequencies would be presented, generally with smaller physical dimensions. The present analysis does not account for the potential simultaneous interference or attentional issues present in a practical online scenario. Another limitation is that the chirp signal, while slowly varying over the selected time windows, is not of a fixed frequency and had a limited duration. Thus there may not have been adequate time to fully entrain certain frequencies, and the CCA may have a slight bias due to the variation in frequency over the time windows. Another consideration is that the selected features may be specific to the CCA approach and may be suboptimal if other SSVEP feature extraction and classification schemes are employed. Through examination of the sensitivity of the ITR and selected features to the near-optimal solutions shown in Figure 5, the selected next-best feature histograms generally have overlapping densities with the best solutions and slight changes in the selected features lead to comparable ITR performances. Nevertheless, it is believed that the present analysis provides a comprehensive evaluation of the most discriminable frequency feature sets for a given number of classes, which should serve as a reference and a guide for standardization of generalized SSVEP stimulus frequencies.

## ACKNOWLEDGMENT

This work was funded in part by the National Science Foundation (1064912).

## REFERENCES

[1] J. R. Wolpaw and E. W. Wolpaw, Eds., *Brain-computer interfaces : principles and practice*. Oxford, New York: Oxford University Press.

[2] M. Middendorf, G. McMillan, G. Calhoun, and K. S. Jones, "Brain-computer interfaces based on the steady-state visual-evoked response." *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, no. 2, pp. 211–4, Jun.

[3] B. Allison and T. Luth, "BCI demographics: How many (and what kinds of) people can use an SSVEP BCI?" *Neural Systems and ...*, no. 2, pp. 107–116.

[4] C. S. Herrmann, "Human EEG responses to 1100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Experimental Brain Research*, no. 3-4, pp. 346–353, Apr.

[5] X. Chen, Z. Chen, S. Gao, and X. Gao, "A high-ITR SSVEP-based BCI speller," *Brain-Computer Interfaces*, no. December, pp. 1–11, Sep.

[6] K. B. Ng, A. P. Bradley, and R. Cunnington, "Stimulus specificity of a steady-state visual-evoked potential-based brain-computer interface." *Journal of neural engineering*, no. 3, p. 036008, Jun.

[7] Y. Zhang, J. Jin, X. Qing, B. Wang, and X. Wang, "LASSO based stimulus frequency recognition model for SSVEP BCIs," *Biomedical Signal Processing and Control*, no. 2, pp. 104–111, Mar.

[8] D. Zhu, J. Bieger, G. Garcia Molina, and R. M. Aarts, "A survey of stimulation methods used in SSVEP-based BCIs." *Computational intelligence and neuroscience*, p. 702357, Jan.

[9] R. Ku, A. Duszyk, P. Milanowski, M. abcki, M. Bierzyska, Z. Radzikowska, M. Michalska, J. Zygierewicz, P. Suffczyski, and P. J. Durka, "On the quantification of ssvep frequency responses in human eeg in realistic bci conditions," *PLoS One*, vol. 8, no. 10, 2013.

[10] A. Duszyk, M. Bierzyska, Z. Radzikowska, P. Milanowski, R. Ku, P. Suffczyski, M. Michalska, M. Labcki, P. Zwoliski, and P. Durka, "Towards an optimization of stimulus parameters for brain-computer interfaces based on steady state visual evoked potentials," *PLoS One*, vol. 9, no. 11, 2014.

[11] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs." *IEEE transactions on bio-medical engineering*, no. 6 Pt 2, pp. 1172–6, Jun.

[12] G. Bin, X. Gao, Y. Wang, and B. Hong, "VEP-based brain-computer interfaces: time, frequency, and code modulations," *Computational*, no. November, pp. 22–26.

[13] N. R. Waytowich and D. J. Krusienski, "Novel Characterization of the Steady-State Visual Evoked Potential Spectrum of EEG," *Braink KDD*, 2014.

[14] T. Tu, Y. Xin, X. Gao, and S. Gao, "Chirp-modulated visual evoked potential as a generalization of steady state visual," vol. 016008, 2012.

[15] F. Sharbrough, C. E. Chatrain, R. P. Lesser, H. Luders, M. Nuwer, and T. W. Picton, "Amearican Electroencephalographic Society guidelines for standard electrode position nomenclature," *J. Clin Neurophysiol*, vol. 8, pp. 200–202, 1991.

[16] G. Schalk, "BCI2000, a general-purpose brain-computer interface (BCI) system," *IEEE Trans. Biomed Eng*, vol. 51, pp. 1034–1043, 2004.

[17] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989.

[18] D. A. Peterson, J. N. Knight, M. J. Kirby, C. W. Anderson, and M. H. Thaut, "Feature selection and blind source separation in an eeg-based brain-computer interface," *EURASIP J. Appl. Signal Process.*, vol. 2005, pp. 3128–3140, Jan. 2005.

[19] S. Koço and C. Capponi, "On multi-class learning through the minimization of the confusion matrix norm," *ArXiv e-prints*, Mar. 2013.

[20] L. Ralaivola, "Confusion-based online learning and a passive-aggressive scheme," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 3284–3292.

[21] S. Baluja and R. Caruana, "Removing the genetics from the standard genetic algorithm," Pittsburgh, PA, USA, Tech. Rep., 1995.