# IMPROVING BCI-BASED COLOR VISION ASSESSMENT USING GAUSSIAN PROCESS REGRESSION

*Hadi Habibzadeh*[1,2], *Kevin J. Long*[2], *Ally E. Atkins*[3,2], *Daphney-Stavroula Zois*[1,2], *James J. S. Norton*[2,1]

[1]Department of Electrical and Computer Engineering, University at Albany, State University of New York, Albany, NY
[2]National Center for Adaptive Neurotechnologies, Stratton VA Medical Center, US Department of Veterans Affairs, Albany, NY
[3]Department of Biological Sciences, Union College, Schenectady, NY

## ABSTRACT

We present metamer identification plus (metaID+), an algorithm that enhances the performance of brain-computer interface (BCI)-based color vision assessment. BCI-based color vision assessment uses steady-state visual evoked potentials (SSVEPs) elicited during a grid search of colors to identify metamers—light sources with different spectral distributions that appear to be the same color. Present BCI-based color vision assessment methods are slow; they require extensive data collection for each color in the grid search to reduce measurement noise. metaID+ suppresses measurement noise using Gaussian process regression (i.e., a covariance function is used to replace each measurement with the weighted sum of all of the measurements). Thus, metaID+ reduces the amount of data required for each measurement. We evaluated metaID+ using data collected from ten participants and compared the sum-of-squared errors (SSE; relative to the average grid of each participant) between our algorithm and metaID (an existing algorithm). metaID+ significantly reduced the SSE. In addition, metaID+ achieved metaID's minimum SSE while using 61.3% less data. By using less data to achieve the same level of error, metaID+ improves the performance of BCI-based color vision assessment.

***Index Terms***— Metamer identification, kernelized learning, grid search, SSVEPs, evoked potentials

## 1. INTRODUCTION

Congenital and acquired color vision deficiencies (CVDs) affect more than 20% of the population [1]. CVDs can be caused by genetics, aging, diseases of the eye or retinal tract, environmental toxins, and some medications [2]; they impair daily life (e.g., driving), limit career options, and affect school performance [3, 4, 5]. Most CVDs are not treatable [6]; however, early screening promotes informed decision making and improves quality of life for affected individuals [7].

The current "gold standard" method for color vision assessment, the anomaloscope, evaluates color vision by identifying *metamers* [8]—light sources with different spectral distributions that appear to be the same color. The anomaloscope for red-green cone deficiencies includes two adjustable light sources: a monochromatic light source consisting of a single primary that emits ∼590 nm (i.e., amber) light and a dichromatic light source consisting of two primaries, each emitting light at ∼670 nm (i.e., red) and ∼546 nm (i.e., green), respectively [2]. During testing, the user manually adjusts the luminances of the two primaries in the dichromatic light source until the dichromatic light source is the same color as the monochromatic light source at a fixed luminance. The ratio of the red to green primaries needed to match the color of the amber primary is different for people with and without CVDs. The anomaloscope has limitations. For example, metamer identification depends on the user's behavioral responses; thus, the anomaloscope is impractical for children, people with motor deficits, those with cognitive impairments, or individuals who are motivated to defeat the test [9, 10].

Norton et al. [11] recently demonstrated a brain-computer interface (BCI) for color vision assessment. This BCI functions similarly to the anomaloscope, but identifies metamers using brain responses instead of behavior. During testing, users are instructed to look at a visual stimulus alternating between a monochromatic and a dichromatic light source while their brain activity is recorded with electroencephalography (EEG). If the two light sources are different colors, the stimulus elicits a steady-state visual evoked potential (SSVEP). Otherwise, the stimulus elicits no SSVEP.

Norton et al.'s [11] BCI uses a grid search of colors to perform metamer identification. The performance of their BCI depends on accurate measurement of SSVEPs. Unfortunately, measurements of SSVEPs are noisy. Norton et al. [11] reduce noise by conducting extensive data collection (∼6 s) for each measurement. This, however, decreases the performance

Corresponding author: Hadi Habibzadeh (hhabibzadeh@albany.edu).
The National Center for Adaptive Neurotechnologies is supported by the National Institute of Biomedical Imaging and Bioengineering of the NIH (P41 EB018783 (JRWolpaw)). This material is the result of works supported by resources at the Stratton VA Medical Center in Albany, NY.

of the BCI. We recently demonstrated metaID [12] to reduce noise in the measurement of SSVEPs using less data. For each measurement, metaID extracts multiple features from the SSVEP, and uses correlations between these features to reduce measurement noise. metaID does not, however, consider correlations between measurements.

Here, we propose metaID+, an enhancement to metaID that uses Gaussian process regression to further reduce measurement noise in BCI-based color vision assessment. metaID+ replaces each measurement with a weighted sum of all other measurements, assigning large weights to correlated measurements (i.e., locations in the grid search with similar colors) and small weights to uncorrelated measurements. We evaluated metaID+ using data collected from ten participants and compared the sum-of-squared errors (SSE; relative to the average grid of each participant) between our algorithm and metaID [12]. metaID+ significantly reduced the SSE. In addition, metaID+ achieved metaID's minimum SSE while using 61.3% less data.

## 2. metaID+

metaID+ uses metaID to obtain SSVEP measurements. Thus, we first provide a brief introduction to metaID.

### 2.1. metaID

metaID [12] uses the filter-bank implementation of CCA (fbCCA) to extract multiple SSVEP features from EEG signals [13]. The distributions of the features are modeled using multi-variate normal (MVN) distributions under two hypotheses: the hypothesis that the colors of the stimulus are metamers ($H_1$) and the hypothesis that the colors of the stimulus are not metamers ($H_0$). metaID uses the result of the log-likelihood ratio test between these two hypotheses as the SSVEP measurement ($y \in \mathbb{R}$). Using MVN distributions enables metaID to extract information from both the amplitudes and correlations of SSVEP features.

### 2.2. metaID+

We use $x_{\text{ref}} \in \mathbb{R}_{\geq 0}$ to denote the fixed luminance of the monochromatic light source and $\mathbf{x} \in \mathbb{R}^2_{\geq 0}$ to denote the luminances of the dichromatic light source. The two-dimensional space of $\mathbf{x}$ can be discretized by restricting $\mathbf{x} \in \mathcal{X}$, where $\mathcal{X}$ is the finite and known set of all possible values for $\mathbf{x}$. metaID+ uses metaID to obtain each measurement for $(x_{\text{ref}}, \mathbf{x})$. We assume each measurement $y(x_{\text{ref}}, \mathbf{x})$ can be expressed as a noisy observation of an unknown noise-free measurement $f(x_{\text{ref}}, \mathbf{x})$:

$$y(x_{\text{ref}}, \mathbf{x}) = f(x_{\text{ref}}, \mathbf{x}) + \epsilon, \tag{1}$$

where $\epsilon$ is independent and identically distributed (iid) noise with mean $\mu = 0$ and variance $\sigma_n^2$ and $y$ and $f \in \mathbb{R}$. metaID+ assumes $f(x_{\text{ref}}, \mathbf{x})$ is a Gaussian process with the mean function $m(\mathbf{x}) = 0$ (i.e., measurements are centered) and the covariance function $k(\mathbf{x}, \mathbf{x}')$:

$$f(\mathbf{x}) \sim \mathcal{GP}\left(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')\right), \tag{2}$$

where the constant $x_{\text{ref}}$ is omitted to reduce notational complexity. The covariance function computes the covariance between every $\mathbf{x}$ and $\mathbf{x}' \in \mathcal{X}$ and is independent of $y$ and $f$. metaID uses an isotropic covariance function, discarding the covariances between similar measurements. However, it is expected that $\mathbf{x}_1 \approx \mathbf{x}_2 \implies f(\mathbf{x}_1) \approx f(\mathbf{x}_2)$. Thus, using a full covariance function can improve the estimation of $f$.

Let $\mathbf{y} \in \mathbb{R}^N$ be the vector of all measurements and $X \in \mathbb{R}^{N \times 2}$ be the corresponding matrix of luminances, where $N = |\mathcal{X}|$ is the cardinality of $\mathcal{X}$. $\epsilon$ is iid, resulting in the following prior on the covariance function of $\mathbf{y}$ ($\Sigma_{\mathbf{y}}$):

$$\Sigma_{\mathbf{y}} = K(X, X) + \sigma_n^2 I, \tag{3}$$

where $K(X, X) \in \mathbb{R}^{N \times N}$ is the covariance function of $X$ and $I \in \mathbb{R}^{N \times N}$ is the identify matrix. Thus, the joint prior distribution of $\mathbf{y}$ and our estimates of the noise-free measurements $\hat{\mathbf{f}}$ can be written as:

$$\begin{bmatrix} \mathbf{y} \\ \hat{\mathbf{f}} \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K + \sigma_n^2 I & K \\ K & K \end{bmatrix}\right), \tag{4}$$

where we replaced $K(X, X)$ with $K$ to simplify the notation. Conditioning the prior on the observed measurements, we obtain the following posterior:

$$\begin{aligned} \hat{\mathbf{f}} | X, \mathbf{y} &\sim \mathcal{N}\left(\mathbb{E}\left[\hat{\mathbf{f}} \middle| X, \mathbf{y}\right], \Sigma_{\hat{\mathbf{f}}}\right), \\ \mathbb{E}\left[\hat{\mathbf{f}} \middle| X, \mathbf{y}\right] &= K[K + \sigma_n^2 I]^{-1}\mathbf{y}, \\ \Sigma_{\hat{\mathbf{f}}} &= K - K[K + \sigma_n^2 I]^{-1}K. \end{aligned} \tag{5}$$

The expected value in Eq. (5) replaces each noisy measurement with the weighted sum of the other measurements; the weight matrix, $W \triangleq K[K + \sigma_n^2 I]^{-1}$, assigns large weights to more correlated measurements and small weights to less correlated measurements.

Algorithm 1 shows the implementation of metaID+. Line 1 collects EEG signals for each luminance in $X$ and uses metaID to convert the signals to SSVEP measurements. The sensing budget $\lambda$ determines the amount of EEG samples collected at each luminance in $X$. Line 2 centers the collected measurements. Line 3 and 4 use the Cholesky decomposition and backward and forward substitution (represented by $\backslash$ symbol) to efficiently solve Eq. (5) [14, 15].

## 3. METHOD

### 3.1. Participants

To test metaID+ we collected data from 10 participants (S01–S10) who had no history of neurological disorders or CVDs. The Stratton VAMC IRB approved all the experiments.

**Algorithm 1** metaID+

**Input:** sensing budget $\lambda$, dichromatic luminances $X$, metamer identification algorithm `metaID`, noise variance $\sigma_n^2$, covariance function `K`

**Output:** estimates of noise-free measurements $\hat{\mathbf{f}}$

1: $\mathbf{y} =$ collect/pre-process data using metaID, $\lambda$, and $X$
2: $\mathbf{y} = \mathbf{y}$ - `mean` $(\mathbf{y})$
3: $\mathtt{LK} = \mathtt{cholesky}\,(\mathtt{K} + \sigma_n^2 \mathbf{I})$
4: $\hat{\mathbf{f}} = \mathtt{K} \times (\mathtt{LK}^{\mathtt{T}} \backslash (\mathtt{LK} \backslash \mathbf{y}))$

### 3.2. Stimulator

The stimulator used in our experiments was described in [11]; it consists of two light sources, a monochromatic amber (590 nm) light source generated by a single LED and a dichromatic red (625 nm) and green (525 nm) light source generated by two LEDs. All three LEDs (amber, red, and green) were physically located on the same die. Light from each LED passed through a lens before being diffused to generate spatially uniform light. The luminance of each LED was controlled using 10-bit PWM from a Teensy microcontroller. Thus, there were 1024 different luminances for each LED; a luminance of 0 denoted when the LED was off and a luminance of 1023 denoted when the LED was at its maximum luminance. We defined $x_{\text{ref}}$ as the PWM setting for the amber LED and $\mathbf{x}$ as the PWM settings for the red and green LEDs.

### 3.3. Experiments

In all experiments, the luminance of the monochromatic source was fixed at $x_{\text{ref}} = 600$, while the luminances of the dichromatic light source were selected from the grid search, $\mathcal{X}$, defined as: $\mathcal{X} \triangleq \Big\{\{0, 24, ..., 312\} \times \{0, 12, ..., 132\}\Big\}$, where $\times$ denotes the Cartesian product. During each data collection *session*, participants completed one measurement for each $\mathbf{x} \in \mathcal{X}$. Each measurement included a six-second stimulation period and a one-second inter-stimulus interval (ISI). During each stimulation period, the stimulator alternated between the two light sources every 50 ms (i.e., flashed at 10 Hz). S3 and S8 completed seven and three sessions, respectively. All other participants completed five sessions.

### 3.4. Data Collection and Analysis

Sixteen channels of EEG data (O1, Oz, O2, PO7, PO3, POz, PO4, PO8, P5, P3, P1, Pz, P2, P4, P6, and CPz; referenced to Cz; ground electrode on the left ear lobe; see [16]) were collected using a g.USB amplifier. Data were sampled at 256 samples per second and saved using BCI2000 [17].

During analysis, all data were filtered using a 12th order IIR bandpass filter with zero-phase shift and cutoff frequencies of 5 Hz and 55 Hz. We used fbCCA to extract SSVEPs features from each measurement [18, 13]. Our implementation of fbCCA used four harmonics and eight subbands (with cutoff frequencies of $i \times 5$ Hz and 55 Hz for $i \in \{1, 2, ..., 8\}$).

**Algorithm 2** Training algorithm for metaID+.

**Input:** the ground truth for each participant $\mathbf{f}_{\text{truth}}$, preprocessed and augmented SSVEP measurements $\mathbf{y}_{\mathtt{train}}$

**Output:** kernel function `K` and noise variance $\sigma_n^2$

1: $\mathtt{K} =$ average of all empirical covariances in $\mathbf{y}_{\mathtt{train}}$
2: **for** for each $\beta \in [0.25, 0.5, ..., 32]$ **do**
3: $\quad\sigma_n^2 = \beta \times \mathtt{max}(\mathtt{K})$
4: $\quad$**for** each participant $p$ **do**
5: $\quad\quad\mathbf{y} = \mathbf{y}_{\mathtt{train}}[p]$
6: $\quad\quad\hat{\mathbf{f}} =$ apply lines 2–4 from Algorithm 1 to $\mathbf{y}$
7: $\quad\quad\mathtt{cost} = \mathtt{COV}(\mathbf{f}_{\text{truth}}, \hat{\mathbf{f}})$
8: $\quad$**end for**
9: **end for**
10: pick $\sigma_n^2 = \beta \times \sigma^2$ that minimized the cost

Under $H_0$, the mean vector $\mu_0$ and covariance matrix $\Sigma_0$ of the MVN model for metaID were estimated using the log of 10 Hz SSVEP features. Under $H_1$, the mean vector $\mu_1$ and covariance matrix $\Sigma_1$ were estimated based on the average of 9.5 Hz and 10.5 Hz non-SSVEP features. To model non-SSVEPs, we used $\mu_1 = \alpha\mu_{\text{non-SSVEP}}$ and $\Sigma_1 = \sigma^2_{\text{non-SSVEP}} I$, where $\alpha$ was an arbitrary small number (i.e., $\alpha = 0.001$), $\mu_{\text{non-SSVEP}}$ was the empirical mean of log of non-SSVEP features, $\sigma^2_{\text{non-SSVEP}}$ was the largest element in the empirical covariance matrix of log of non-SSVEP features, and $I$ was the identity matrix of proper size. Using a small $\alpha$ (as opposed to $\alpha = 0$) improved the algorithm's numerical stability.

We used Algorithm 2 to train metaID+. The algorithm requires the pre-processed and augmented measurements $\mathbf{y}_{\mathtt{train}}$. To obtain $\mathbf{y}_{\mathtt{train}}$, we collected EEG signals at each luminance in $X$ using a sensing budget of six seconds. The training sensing budget was larger than the test sensing budget. Thus, the preprocessing step involved an overlapping sliding window to break the EEG signals into $\lambda$-second segments. For each segment, participant, and location in $X$, we obtain a measurement vector $\mathbf{y}$ using metaID. Furthermore, we augmented the training dataset by adding shifted replicas of each $\mathbf{y}$. We converted $\mathbf{y}$ into two-dimensional grid (with 14 unique red luminances on the y-axis and 12 unique green luminances on the x-axis), used `numpy's roll` function to shift the rows and columns, and flattened the results to get the shifted replica of $\mathbf{y}$.

For each participant, each $\lambda$-second long segment, and each shifted replica of measurement vectors in $\mathbf{y}_{\mathtt{train}}$, we computed the empirical covariance matrix and used the average of all these covariance matrices as the estimate for `K`. The largest element in `K` was selected as the baseline of the noise variance. In lines 3–11 of Algorithm 2, we scale the noise variance using the scalar $\beta$, compute the de-noised grid $\hat{\mathbf{f}}$, and measure its difference from the ground truth. Finally, we pick the $\beta$ that gave the best results. Both metaID and

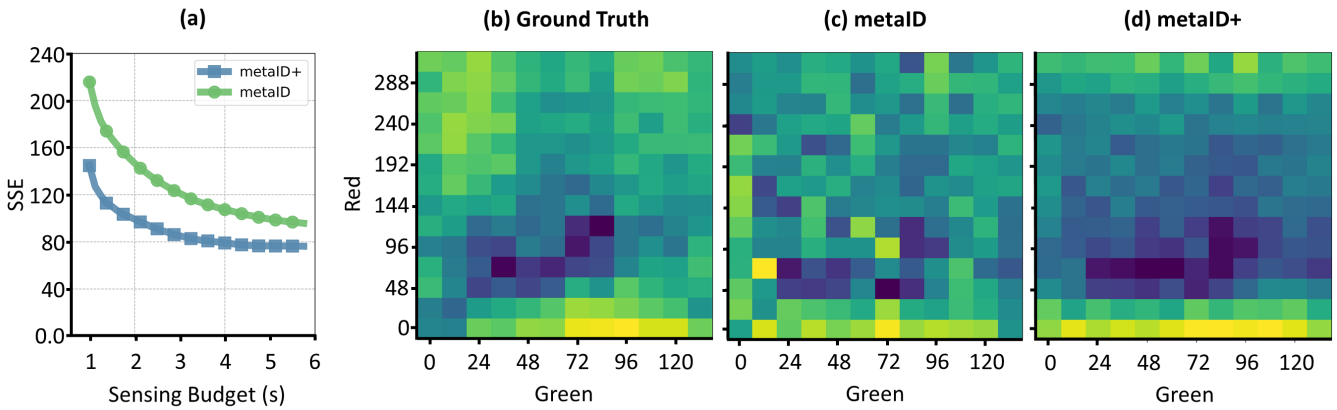


**Fig. 1**. (a) Average sum of squared errors (SSE) for metaID and metaID+ for different sensing budgets (i.e., $\lambda$; in seconds). (b–d) Example grids for the ground truth, metaID, and metaID+. The 'red' and 'green' axes represent luminances of the red and green LEDs, respectfully. The colors at each location in the grid represent the size of the measurements. (b) Ground truth for S05. (c) metaID for S05–session two ($\lambda = 1$, SSE=198.3). (d) metaID+ for S05–session two ($\lambda = 1$, SSE=100.6).

metaID+ were trained using cross-validation; in each round, the data from one participant were reserved for evaluation and data from other nine participants were used for training.

### 3.5. Performance Metric

To get the ground truth for each participant, we used metaID with the maximum sensing budget (i.e., $\lambda = 5.625$ s) to obtain all measurements in the grid search for each session and then averaged the grid searches across sessions. Performance of metaID and metaID+ was evaluated using SSE between the ground truth and the output of each algorithm.

## 4. NUMERICAL RESULTS

Figure 1a compares the SSE (averaged across all participants) of metaID+ and metaID for different $\lambda$'s. metaID+ reduced SSE for all sensing budgets ($p < 0.001$, using one-sided paired t-test). metaID had a minimum SSE (97.8) at the average sensing budget of 5.5 s. metaID+ achieved a similar SSE (96.4) using an average sensing budget of 2.1 s. In other words, metaID+ achieved metaID's minimum SSE while using 61.3% less data. Figures 1b, 1c, and 1d show example measurement grids for the ground truth, metaID, and metaID+, respectively. The SSE for metaID+ in Fig. 1c is $= 100.6$, whereas the SSE for metaID in Fig. 1d is $= 198.3$.

## 5. DISCUSSION

metaID+ significantly reduced the measurement error of BCI-based color vision assessment compared to metaID, especially for smaller sensing budgets. In addition, metaID+ achieved metaID's minimum SSE using 61.3% less data. By using less data to achieve the same error, metaID+ improves the performance of BCI-based color vision assessment.

metaID+ uses Gaussian process regression, which is computationally complex due to the Cholesky decomposition ($\mathcal{O}(N^3)$). In BCI-based color vision assessment, $N$ is on the order of 100 (in our paper, $N = 168$). Using a standard desktop computer, our algorithm took $\sim 0.03$ s to compute (Intel i7 9700). Thus, the computational complexity of Gaussian process regression is not a concern for metaID+ even for online use.

metaID+ is easy to train. Algorithm 2 showed how metaID+ can be estimated from data, without using covariance function kernels (e.g., RBF, Matern kernel [19]). Color vision is complex and non-linear. Thus, a linear combination of multiple kernels is needed to model the correlations among measurements. However, using multiple kernels (and deciding what kernels to use) requires expertise on color vision, increases the number of hyper-parameters, and complicates training. In contrast, Algorithm 2 can model complexities of the grid search using a single hyper-parameter (i.e., $\beta_\lambda$).

Future work on metaID+ should consider its implementation for online use. In addition, while our results showed that metaID+ improves BCI-based color vision assessment, further experiments are needed to demonstrate metaID+'s ability to identify congenital and acquired CVDs.

## 6. CONCLUSION

We presented metaID+, an algorithm that uses Gaussian process regression to reduce measurement noise during BCI-based color vision assessment. We compared metaID+ to metaID (an existing algorithm) using experimental data from 10 participants. Results showed that metaID+ achieved metaID's minimum SSE while using 61.3% less data, greatly enhancing the performance of BCI-based color vision assessment.

## 7. REFERENCES


[1] Marilyn E Schneck, Gunilla Haegerstrom-Portnoy, Lori A Lott, and John A Brabyn, "Comparison of panel D-15 tests in a large older population," *Optometry and Vision Science: Official Publication of the American Academy of Optometry*, vol. 91, no. 3, pp. 284, 2014.

[2] Matthew P Simunovic, "Acquired color vision deficiency," *Survey of Ophthalmology*, vol. 61, no. 2, pp. 132–155, 2016.

[3] Brian Webster, "Colour vision deficiency: The 'unseen' disability," *British Journal of Nursing*, vol. 30, no. 8, pp. 468–469, 2021.

[4] Walter T Delpero, Hugh O'Neill, Evanne Casson, and Jeff Hovis, "Aviation-relevent epidemiology of color vision deficiency," *Aviation, Space, and Environmental Medicine*, vol. 76, no. 2, pp. 127–133, 2005.

[5] Antonio Tagarelli, Anna Piro, Giuseppe Tagarelli, Pasquale Bruno Lantieri, Domenico Risso, and Rosario Luciano Olivieri, "Colour blindness in everyday life and car driving," *Acta Ophthalmologica Scandinavica*, vol. 82, no. 4, pp. 436–442, 2004.

[6] MP Simunovic, "Colour vision deficiency," *Eye*, vol. 24, no. 5, pp. 747–755, 2010.

[7] Barry L Cole, "Assessment of inherited colour vision defects in clinical practice," *Clinical and Experimental Optometry*, vol. 90, no. 3, pp. 157–175, 2007.

[8] Theresa J Squire, Marisa Rodriguez-Carmona, Anthony DB Evans, and John L Barbur, "Color vision tests for aviation: Comparison of the anomaloscope and three lantern types," *Aviation, Space, and Environmental Medicine*, vol. 76, no. 5, pp. 421–429, 2005.

[9] Richard B Lomax, Peter Ridgway, and Maureen Meldrum, "Does occupational exposure to organic solvents affect colour discrimination?," *Toxicological Reviews*, vol. 23, no. 2, pp. 91–121, 2004.

[10] Jason S Ng and Sophia C Liem, "Can the Farnsworth D15 color vision test be defeated through practice?," *Optometry and Vision Science*, vol. 95, no. 5, pp. 452–456, 2018.

[11] James JS Norton, Grace F DiRisio, Jonathan S Carp, Amanda E Norton, Nicholas S Kochan, and Jonathan R Wolpaw, "Brain-computer interface-based assessment of color vision," *Journal of Neural Engineering*, vol. 18, no. 6, pp. 066024, 2021.

[12] Hadi Habibzadeh, Daphney-Stavroula Zois, and James J S Norton, "metaID: A metamer identification algorithm for improving BCI-based color vision assessment," in *Asilomar Conference on Signals, Systems, and Computers*, 2021.

[13] Xiaogang Chen, Yijun Wang, Shangkai Gao, Tzyy-Ping Jung, and Xiaorong Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface," *Journal of Neural Engineering*, vol. 12, no. 4, pp. 046008, 2015.

[14] numpy.org, "Cholesky decomposition," `https://numpy.org/doc/stable/reference/generated/numpy.linalg.cholesky.html`, 2021.

[15] numpy.org, "Least-square solver," `https://numpy.org/doc/stable/reference/generated/numpy.linalg.lstsq.html`, 2021.

[16] GE Chatrian, E Lettich, and PL Nelson, "Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities," *American Journal of EEG Technology*, vol. 25, no. 2, pp. 83–92, 1985.

[17] Gerwin Schalk, Dennis J McFarland, Thilo Hinterberger, Niels Birbaumer, and Jonathan R Wolpaw, "BCI2000: A general-purpose brain-computer interface (BCI) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[18] Zhonglin Lin, Changshui Zhang, Wei Wu, and Xiaorong 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.

[19] Carl Edward Rasmussen, "Gaussian processes in machine learning," in *Summer School on Machine Learning*. Springer, 2003, pp. 63–71.