Capturing Pictures from Human Vision Using SSVEP and Lock-in Amplifier

Danson Evan Garcia, Kai Wen Zheng, Yi (Summer) Tao, Yi Liu, Steve Mann
MannLab Canada, 330 Dundas Street West, Toronto, Ontario, M5T 1G5

Abstract—We present a novel way of using one's eye to capture an image of what it "sees" through the use of steady-state visually-evoked potentials (SSVEP). Existing methods leveraging response patterns for SSVEP visual image reconstruction show lossy reconstruction and have a lengthy scanning process. With our signal acquisition procedure, data collection requirements are significantly decreased while still improving the signal clarity. The data for image reconstruction were collected from the Oz positioned electrode using a low-cost, wearable electroencephalography (EEG) device. For image reconstruction, software-defined lock-in amplifier (LIA) and discrete Fourier transform (DFT) signal processing methods are analyzed and compared.

Index Terms—Signal processing, visual field reconstruction, brain-computer interfaces (BCI), steady-state visually evoked potential (SSVEP), lock-in amplifier, pattern recognition.

I. Background and Introduction

Vision, sensation, and perception have long been important fields of study [1], [2]. The knowledge of how humans perceive and sense the world around them can be applied to create body sensor networks [3], smart prosthetics [4], [5], and implants [6] that help those with severe disabilities and those with difficulties in performing perceptual, sensational, or motor tasks. Sensors [7], wearable computing [6], [8], [9], and brain-computer interface (BCI) [10]–[12] have become increasingly relevant to medicine and health care as early detection and monitoring systems. As these techniques continue to develop, the requirements for our understanding of the capabilities of these devices and human senses will become increasingly significant [13], [14].

Metaveillance is the sensing of sensing [15]. Through understanding and visualization of the senses, we may improve our understanding of the perception of both humans and machines and their respective capabilities [16], [17]. This understanding of human and machine perceptions, and human vision, in particular, can greatly influence the progress of imaging and pattern recognition technology, in much the same way as the anatomy of the human eye influenced the camera, and the understanding of the neuron influenced the design of neural networks.

Neuroimaging devices can provide strong entry points of study for metaveillance applied to humans. However, most investigations using neuroimaging techniques such as EEG, fMRI, etc, have barriers to entry due to the need for specialized and expensive equipment [2], [18]. As an example, a current state-of-the-art visualization technique that leverages deep neural networks with 6000 training images to generate images from human vision uses expensive fMRI devices [19]. Thus, to
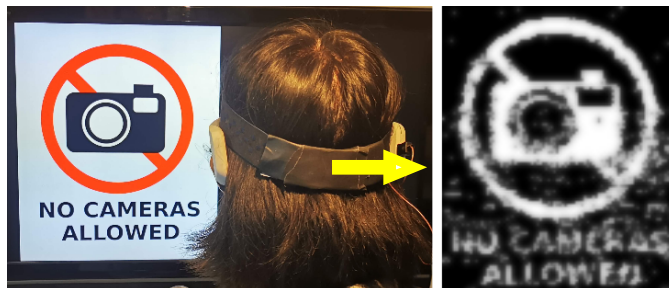

Fig. 1: The conceptual end goal of this research, a user of our system looks at a "NO CAMERAS ALLOWED" sign (left) and captures an image (right) reconstructed from EEG signals using techniques discussed in this paper.

ensure that our project is easily accessible, we have designed our experiments to operate on a low-cost, wearable, 5-channel EEG device. This can potentially shine a new light on how wearable EEG devices are used other than for their existing meditative and mental state monitoring applications [20].

A. SSVEP

Steady-state visually-evoked potentials (SSVEP) are a known technique used in EEG research. SSVEP response patterns are periodic responses from the visual cortex that are evoked through a visual stimulus flickering at specific frequencies. An increase in neural activity at the same frequency of the visual stimulus is present in the attended region [2], [22]. SSVEP has recently engendered growing interests in the scientific community [22]–[26] and is widely used in studies investigating spatial attention [2]. In common practice, SSVEP signals are usually monitored using EEG or functional magnetic resonance imaging (fMRI) [2], [27], and are detected from a large range of frequencies, 1–90 Hz [28], [29]. However, human brains show steady and strong responses only in a limited range of stimulus frequencies.

B. Eye as a Camera

In this paper, we propose several new methods of capturing vision directly from the eye. Previous studies show the possibilities to measure and visualize the ability to see, as well as investigate the augmented reality visualization of human sight [16], [17]. Through the use of SSVEP response patterns, a subject's visual field can then be reconstructed [30].

The prior reconstruction procedure [30] has several drawbacks. The first is the lengthy scanning process, which hinders possible applications. Other concerns include the methods for data acquisition and signal processing techniques, leading to

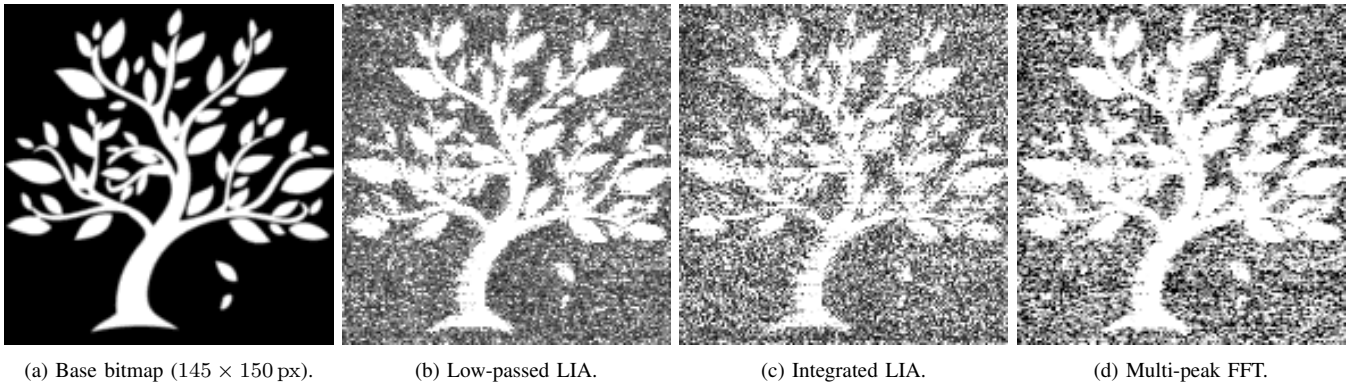| (a) Base bitmap (145×150 px). | (b) Low-passed LIA. | (c) Integrated LIA. | (d) Multi-peak FFT. |

Fig. 2: The images are the visual field reconstructions of the base bitmap of a tree [21] with only bit-value scaling applied to scale the maximum bit value to 255. Each image uses a different reconstruction algorithm. The EEG data are collected by User A with a stimulus frequency of 8.5 Hz. The time duration for each pixel data collection is 1 s for a total experiment time of about 6 h.

significant loss of image reconstruction quality. Based on these drawbacks of the original approach, we hypothesize that by altering the experiment methods and by applying different signal processing algorithms on the raw EEG data, the quality of the reconstructed images will increase. Decreasing the total experimental time is also another requirement to improve the overall user experience and to realize the potential for real-life applications.

Fig. 1 depicts the end goal for our procedure, the concept that human vision can directly serve as a medium for recording visual information.

II. METHODS

The Muse Meditation Headband by InterAxon Inc. is the EEG sensing device used during the experiments. When worn, an external electrode is added and placed at the occipital lobe (Oz) position to detect SSVEP responses. Following the 10-20 system of EEG placement, the Oz position is the most consistent for measuring SSVEP responses at varying frequencies [31].

For the following experimental procedure, we chose 5 different images. The star, panda, and spring tree images showcase our procedure's efficacy on images of increasing geometric complexity, while the "NO CAMERAS ALLOWED" sign and the face portrait represent typical images that one may encounter in their day-to-day life. All images are black and white to simplify the process and ensure the presence of distinct SSVEP responses [30].

A. Experimental Procedure

First, new participants perform a 5-minute frequency detection task using a flickering stimulus window of a black image and a white image of varying frequencies from 5 Hz to 20 Hz at 0.5 Hz increments to discover their optimal SSVEP response frequency. For each target frequency, the stimulus flashes for 6 s with 4 s rest in between frequencies. Fig. 3 shows graphs of raw EEG data and the corresponding outputs from the LIA and FFT algorithms at the target frequency. The FFT uses a 2-second data window, i.e. 512 samples at 256 Hz sampling frequency.

By comparing and contrasting, we select the frequency that elicits the most stable and observable SSVEP responses with the shortest response delay for the following scanning experiments. The response delay is the delay between the stimulus on screen and the response activation in the brain. From Fig. 3a, the LIA and FFT graphs show a steady activation from 2 to 6 s. The LIA graph remains consistently greater than 10, while the FFT slowly ramps up from 2 to 4 s and reaches a steady-state activation. A sharp linear deactivation response to the stimulus follows at 6 s for both methods of analysis. The mild bumps in the LIA graph between 3 to 4 s and 5 to 6 s may represent the user's increased visual attention. Fig. 3b is a sub-optimal SSVEP response graph with an activation that is unstable and a deactivation that is difficult to notice. Unlike in Fig. 3a, the LIA graph in ig. 3b does not consistently reach high activation between 2 to 6 s. The FFT also only shows what may perhaps be a weak activation between 3 to 4 s as well as 4 to 6 s.

Stable SSVEP response (i.e. continuous and distinct activation response) will allow better signal reconstruction quality, and shorter response time will reduce interference and corruption when the stimulus window switches. Once we find a suitable frequency, redoing the frequency detection task is not necessary since the main objective is to find a single frequency that fits a particular person.

Once a stimulus frequency is selected, we select an image for the scanning process. Before slicing the image into smaller sub-images of equal frame size, the selected image is scaled and padded with additional black pixels. The participant may apply side paddings on the image to avoid stimulus flickering between the end of a previous row and the start of the next one. The top and bottom paddings can help capture top and bottom image details easier. A step size defines the vertical and horizontal shift between each sub-image. The step-size and the scaling of the original image determine the resolution of the resulting bitmap image. To illustrate, we slice a panda image into 200×200 px frame images with a step-size of 40 px in Fig. 4. Fig. 4 also shows 2x4 sample sub-images around a point of interest marked in red. We then display

(a) 8.5 Hz frequency test graphs from User A.

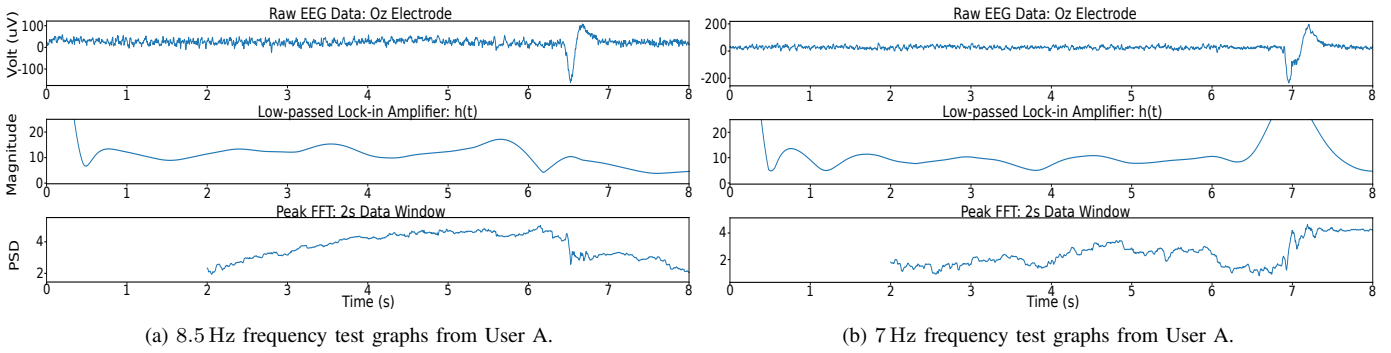(b) 7 Hz frequency test graphs from User A.

Fig. 3: a) and b) are the resulting graph from a frequency test run at 8.5 Hz and 7 Hz respectively for User A. The top graph is the raw EEG data, the middle is magnitude output using Low-passed LIA, and the bottom is the FFT power spectral density (PSD) with the noise floor subtracted. Both processing methods use a 2-second data capture window.
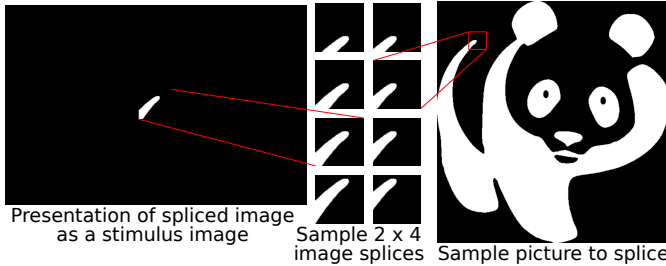

Fig. 4: Panda image [32] (2000×2092 px) is sliced into 200×200 px at 40 px step sizes and the corresponding presentation as a stimulus. The corresponding bitmap is an image with 50×53 px.

these sub-images individually as a stimulus image at the center of a black background application window flickering at the stimulus frequency. Participants focus their eyes at a 3 px-tall indicator located at the center of the application window. The indicator aids in keeping the attention of the participant on the flickering stimulus. After a fixed time duration, the stimulus image moves from left to right and top to bottom, which simulates the picture going across the screen, similar to the regular reading direction.
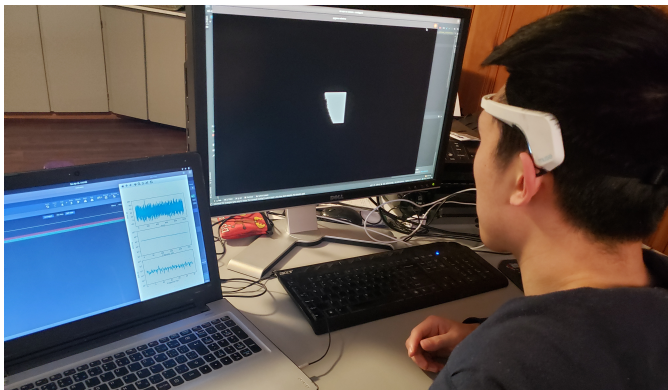

Fig. 5: The experiment proper with the stimulus image presented on the monitor and EEG data recordings from Muse headset to the laptop computer.

Fig. 5 shows a picture of a subject performing the experiment. While the user focuses on the stimulus image, we collect and timestamp the user's EEG data coming from the Muse. The experiment repeats until all the rows are taken.

Throughout the experimental process, the participants can relax and blink normally. Though we encourage participants to be focused, short moments of distraction will not overly influence the final image.

Raw data collected are processed using Fast Fourier Transform (FFT) and the two implementations of software-defined Lock-in Amplifier (LIA) algorithms. We extract the bitmap of a particular pixel from the raw data using the start time and the end time recorded for each stimulus image, while each stimulus image corresponds to one pixel in the final reconstruction.

Fig. 2 shows the corresponding image reconstructions of a tree using each algorithm. These reconstructed images have the highest pixel resolution (145×150 px) among all the images we have taken using the eye as a camera. The original image size is 6000×6000 px. Fig. 8 shows the corresponding image reconstructions of a human face using each algorithm. The original photograph size is 2051×2736 px while the reconstructed images and the target bitmap are 82×102 px.

B. Fast Fourier Transform

Fast Fourier transform [33] is an algorithm that computes the discrete Fourier transform (DFT) [34] of a sequence of data points. The algorithm transforms the signal from its original time-domain representation to components in the frequency domain. Fig. 6 shows a sample graph taken using the algorithm with the first and second SSVEP response harmonic frequencies encircled for emphasis.
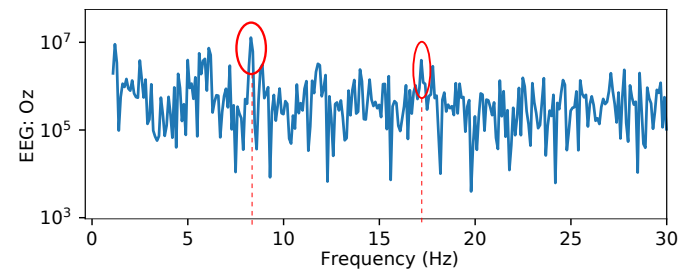

Fig. 6: Sample PSD of EEG data at Oz position from 1-30 Hz. Encircled are the SSVEP response peaks at the first (8.5 Hz) and the second (17 Hz) harmonics. The data window size used in this transformation is 2560 samples.

Prior study [30] examines activation at the target frequency to estimate the strength of the SSVEP responses while considering only the peak activation for a particular pixel, Maximum Peak FFT. However, activations for SSVEP are often not located only at the target frequency but in a triangle of activation surrounding the target frequency, as shown in Fig. 6. Multi-peak FFT, therefore, takes all the activation in a range around the target value and sums them up to receive its bitmap value. We note that FFT requires a minimum number of sampled points for its transformation. With too few sample points computed, the desired frequency may not exist in the output. Therefore, for time windows below $2\,\mathrm{s}$, the FFT data window must be widened artificially in a range around the stimulus duration for data to be processed.
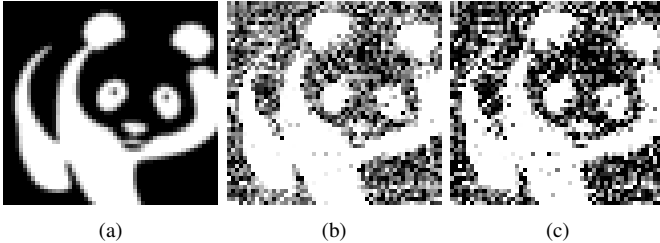

(a)   (b)   (c)

Fig. 7: The EEG data for reconstructing the panda image [32] are taken by User B at $8.5\,\mathrm{Hz}$ with $2\,\mathrm{s}$ pixel time duration. a) Base bitmap image ($53 \times 50\,\mathrm{px}$). b) FFT reconstruction considering the peak FFT at the SSVEP stimulus frequency, Maximum Peak FFT. c) FFT reconstruction considering multiple peaks for a particular SSVEP stimulus frequency, Multi-peak FFT.

TABLE I: Accuracy measurements in choosing between FFT implementation shown in Fig. 7b (Maximum Peak FFT) and Fig. 7c (Multi-peak FFT).

	Maximum Peak FFT	Multi-peak FFT
RMS Accuracy (Acc.)	0.588	0.625
True Positive Rate	0.983	0.964
True Negative Rate	0.550	0.670

Two implementations of FFT, Maximum Peak FFT, and Multi-peak FFT, are used to compute the reconstruction. Fig. 7 shows the difference between the two FFT reconstruction methods after subtracting the noise floor. To limit the algorithms discussed, we consider only the FFT implementation that results in better accuracy. Table I shows the accuracy comparison between the two FFT implementations. We obtain the metrics by comparing the image reconstructions with the base bitmap image in Fig. 7a. RMS accuracy refers to root-mean-square accuracy which derives from $1 - \mathrm{RMSE}$, where RMSE is the root-mean-square error between the reconstruction and the target bitmap. In this paper, the true positive rate (TPR) refers to the percent of correct predictions of white pixels based on the target bitmap while the true negative rate (TNR) refers to the percent of correct predictions for black pixels. As shown, the Multi-peak FFT method performs better than the Maximum Peak FFT method. Although the TPR for Multi-peak FFT is less, when we consider the sum of TPR and TNR, Multi-peak FFT gets a better reconstruction.

C. Lock-in Amplifier

A lock-in amplifier is an analog device that selects and amplifies a particular oscillatory frequency while rejecting all other frequencies [35], [36]. Due to this nature, LIA can extract a signal in an extremely noisy environment as long as this reference signal frequency is precise and known. An advanced form of LIA uses a second reference signal, $90°$ phase-shifted to the first reference. Given a target frequency, two reference signals, $y(i\omega t)$ and $z(i\omega t)$ in Eqn. 1, are multiplied to the original signal. Software-based LIA is useful for our experiment since we want to extract the SSVEP response frequency, and we know the target frequency, f_s, and the corresponding i harmonics. In this paper, we consider only the data obtained from the fundamental frequency. Thus, we introduce two software-based LIA implementations.

$$y(i\omega t) = z(i\omega t + \frac{\pi}{2}), t = (t_1, t_2, \ldots, t_n)$$
$$\omega = 2\pi f_s \tag{1}$$

1) Integrated Lock-in Amplifier: This software-defined LIA uses the integration of sinusoidal functions to cancel out all other frequencies and extract the original signal. As shown in Eqn. 2, the raw EEG data $u(t)$ is multiplied by the reference wave then integrated over n points where n is the total number of data points sampled over the given stimulus duration. The same is done with the reference wave replaced by its $90°$ phase-shifted counterpart.

$$p = \int u(t)y(i\omega t)dt = \frac{1}{n}\sum_{j=1}^{n} u(t_j)y(i\omega t_j)$$
$$q = \int u(t)z(i\omega t)dt = \frac{1}{n}\sum_{j=1}^{n} u(t_j)z(i\omega t_j) \tag{2}$$

The magnitude obtained from p and q removes phase dependence. After scaling by a constant c, the result is the final pixel value in Eqn. 3.

$$pixel = c\sqrt{q^2 + p^2} \tag{3}$$

2) Low-passed Lock-in Amplifier: This software-defined LIA implementation mirrors how a hardware-based LIA performs. Eqn. 4 details the mathematical formula of this implementation. The raw EEG data is first multiplied by the reference wave and its $90°$ phase-shifted counterpart at the target frequency for each pixel frame. Then, a second-order Butterworth low pass filter at $0.7\,\mathrm{Hz}$ is applied to the modified signals. The filtering is necessary in order to yield an approximately time-independent output [37]. The magnitude, α, is obtained from these time-independent outputs.

$$V_y(i\omega t) = lowpass(u(t)y(i\omega t), 0.7\,\mathrm{Hz})$$
$$V_z(i\omega t) = lowpass(u(t)z(i\omega t), 0.7\,\mathrm{Hz})$$
$$\alpha(i\omega t) = \sqrt{V_y(i\omega t)^2 + V_z(i\omega t)^2} \tag{4}$$

Eqn. 5 simply averages the summed magnitudes of the base

(a) Base bitmap (82 × 103 px).   (b) Low-passed LIA.   (c) Integrated LIA.   (d) Multi-peak FFT.
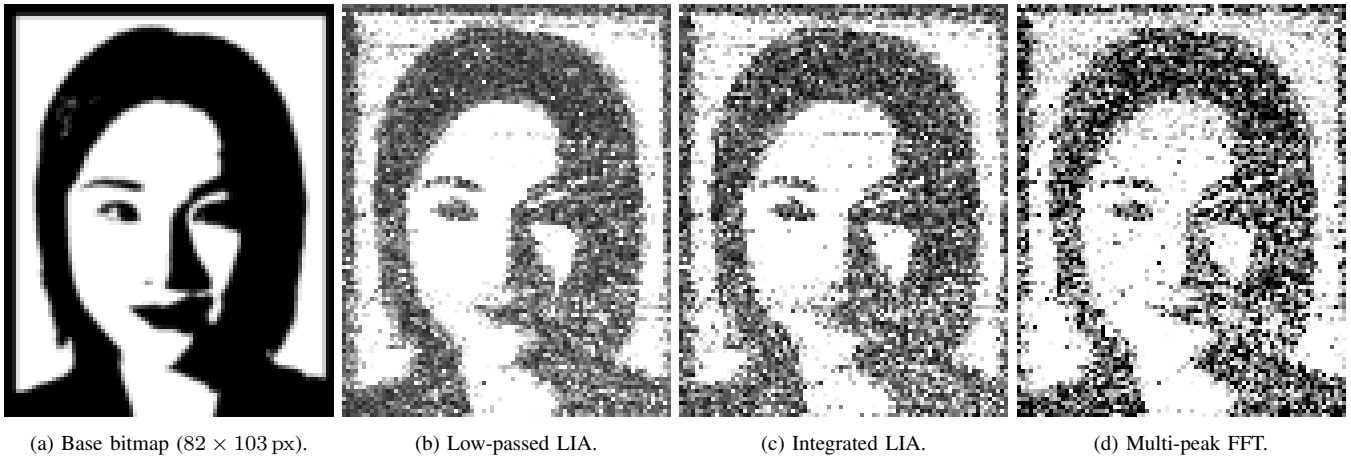
Fig. 8: Portrait image reconstruction using SSVEP response data from multiple users. First 34 px rows are from User A. Next 35 px rows are from User B. Last 34 px rows are from User C. EEG data from User A and B are collected with a stimulus frequency of 8.5 Hz while User C are collected with 8 Hz. Time duration of 1.5 s for each image frame is used by all users.
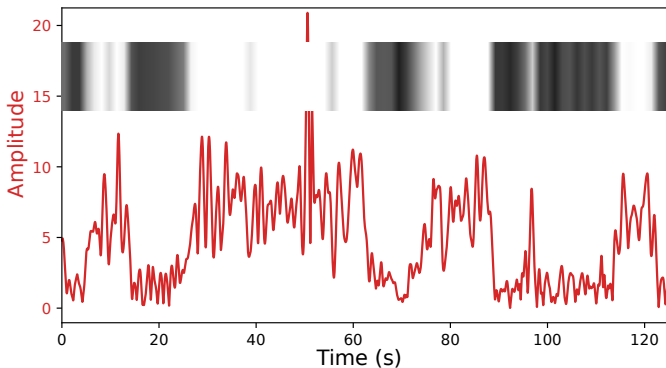

Fig. 9: Example pixel row conversion using the Low-passed LIA. This bitmap corresponds to the 62nd row in Fig. 8b.

frequency at the time range, $t_a, \ldots, t_b$, that corresponds to a particular sub-image frame duration. c is a constant applied to each value found for one entire image bitmap. Hence, $pixel$ is used to indicate the pixel value of a particular sub-image frame.

$$pixel = \frac{c \sum_{j=a}^{b} \alpha(t_j)}{b - a} \tag{5}$$

Fig. 9 illustrates this pixel transformation. The graph shows the $\alpha(t)$ for a particular row of Fig. 8b. The corresponding grayscale 82 px image row generated from SSVEP is overlayed on the graph to match the amplitude pattern.

III. RESULTS AND DISCUSSION

To ensure the the work is reproducible, we obtained results from three users, User A, User B and User C. User A and User B, both male, have their best SSVEP response at 8.5 Hz. The best SSVEP response frequency for User C, female, is at 8 Hz. All three users used the Muse headband device, and the experiments were successfully carried out both together with supervision and individually at different locations.
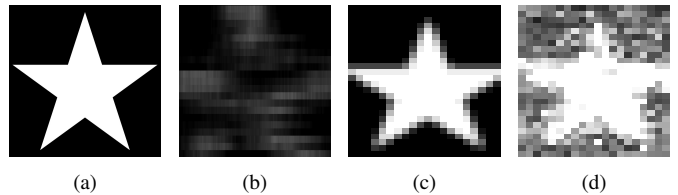

Fig. 10: a) The original star image (1000 × 999 px). b) Reconstruction from the prior art [30] by User C, which takes about 27 min of experiment time. c) The base bitmap of the original image (20 × 26 px). d) Reconstruction using Low-passed LIA at 2 s stimulus duration by User C, which takes about 17 min of experiment time.

A. Capture Image Accuracy and Improvements

Fig. 10 shows the dramatic improvements in the quality of the reconstruction and the experiment time requirements using the methods described earlier. Even though total experiment time is just 17 min, Fig. 10d is clearly a better reconstruction compared to Fig. 10b, which took 27 min of data collection.
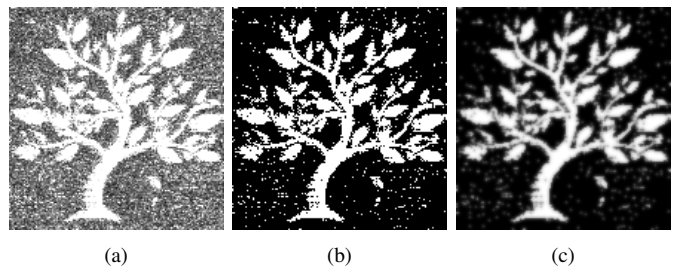

Fig. 11: Denoising the reconstructed image of the tree shown in Fig. 2b. a) Untreated image. b) Result after setting pixel threshold set at 0.92. c) Result after applying total variation denoising [38] set at 0.49 on b).

Fig. 11 further illustrates the merits of our breakthroughs. To elaborate, the most dramatic increase in RMS accuracy came from applying a simple pixel threshold adjustments to classify a black and a white pixel. In Fig. 12, as the pixel threshold increases, the TNR increased substantially compared to the decrease in TPR. We obtained the highest TPR and TNR combination when we set the threshold at the 92nd percentile.
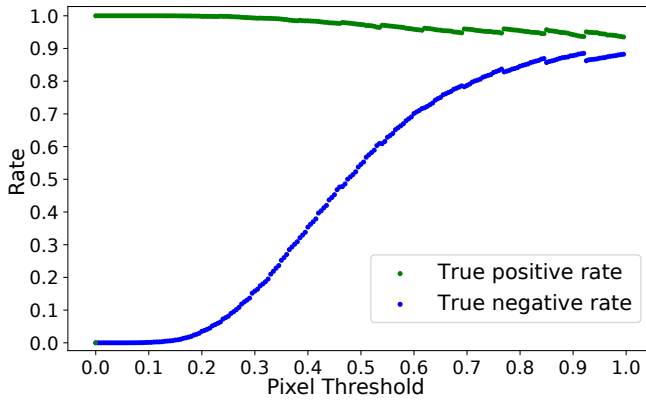
Fig. 12: Showing TPR and TNR with different thresholds set on Fig. 11a. The threshold is set to maximize TPR + TNR. In this case the best threshold is 0.92.

Therefore, whites were data values from the 92nd percentile and above, while the rest were blacks. After thresholding, we applied a total variation denoising [38] algorithm to the image. The best weight found to maximize RMS accuracy is 0.42, resulting in a 0.840 RMS accuracy image. By using simple and accessible denoising techniques, we managed to improve the accuracy of the reconstruction for a feature-rich image from 0.544 to 0.840 based on Table II.

TABLE II: Accuracy measurements for different pictures of interest.

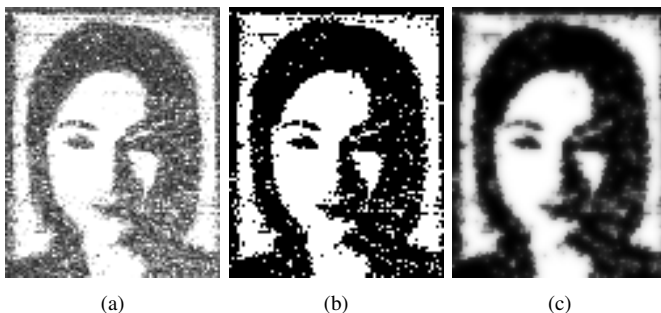|	RMS Acc.	TPR	TNR
Fig. 11a	0.544	0.973	0.546
Fig. 11b	0.731	0.843	0.931
Fig. 11c	0.840	0.869	0.959
Fig. 13a	0.661	0.961	0.704
Fig. 13b	0.703	0.863	0.924
Fig. 13c	0.812	0.881	0.970

(a) (b) (c)

Fig. 13: Denoising the reconstructed image of the portrait shown in Fig. 8b. a) Untreated image. b) Result after setting pixel threshold set at 0.75. c) Result after applying total variation denoising [38] set at 0.34 on b).

Fig. 13 is the same set of denoising algorithms applied on the portrait image shown in Fig. 8b. The accuracy also improves, showing that the process is reproducible. The major contribution of thresholding is to improve the true negative rate while preserving a reasonable true positive rate. Thus, the improvement is greater on bitmaps with more black pixels (Fig. 11) than bitmaps with less black pixels (Fig. 13). The

total variation denoising algorithm removes outlier pixels caused by remaining noises, improving overall RMS accuracy.

B. Stimulus Duration

The effect of stimulus duration on visual reconstruction was examined. Experimental data for different stimulus frame duration were gathered from each user, starting at 0.25 s. The image reconstruction results for a star with increasing stimulus frame duration is shown in Fig. 14 for Low-passed LIA, Fig. 15 for Integrated LIA, and Fig. 16 for Multi-peak FFT. From the results, Multi-peak FFT has the longest stimulus duration required to get any recognizable result, as both 0.25 s, and 0.5 s stimulus duration do not yield any identifiable features of the star. Low-passed LIA is able to retrieve the star at 0.5 s time duration, while Integrated LIA is able to present useful shape information at the stimulus duration of 0.25 s. As a reference, a 0.25 s duration only provides two stimulus periods for either 8 Hz or 8.5 Hz stimulus frequencies.
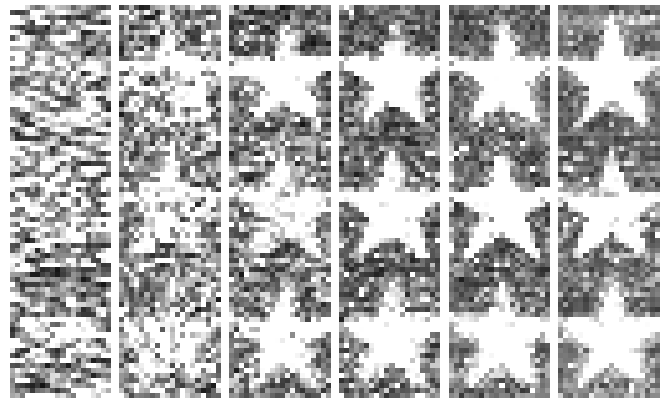

Fig. 14: Low-passed LIA reconstructions of a star at different frame duration by three users. Top row are from User A. Middle row are from User B. Last row are from User C. Stimulus frame duration are increasing from left to right: 0.25 s, 0.5 s, 0.75 s, 1.0 s, 2.0 s and 4.0 s.

Fig. 15: Integrated LIA reconstructions of a star at different frame duration by three users. The top row is from User A. The middle row is from User B. The last row is from User C. Stimulus frame duration are increasing from left to right: 0.25 s, 0.5 s, 0.75 s, 1.0 s, 2.0 s and 4.0 s.

Fig. 17 shows a plot of the RMS accuracy for different stimulus duration taken of the star by each user using the
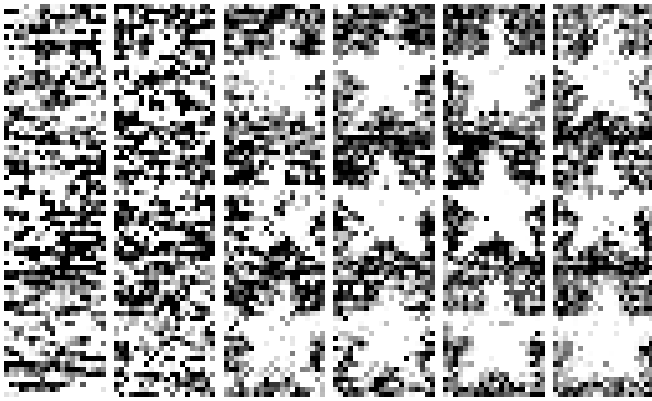
Fig. 16: Multi-peak FFT reconstructions of a star at different frame duration by three users. The top row is from User A. The middle row is from User B. The last row is from User C. Stimulus frame duration are increasing from left to right: 0.25 s, 0.5 s, 0.75 s, 1.0 s, 2.0 s and 4.0 s.

Low-passed LIA method. Though different users have different accuracy curves in relation to the stimulus duration, accuracy clearly trends upward and reaches a steady-state. Between 0 to 2 s, accuracy rises steeply. Afterward, although we can see some improvements, the gains are not as dramatic. The experiment shows that the steady-state accuracy is different for every user albeit only slightly.
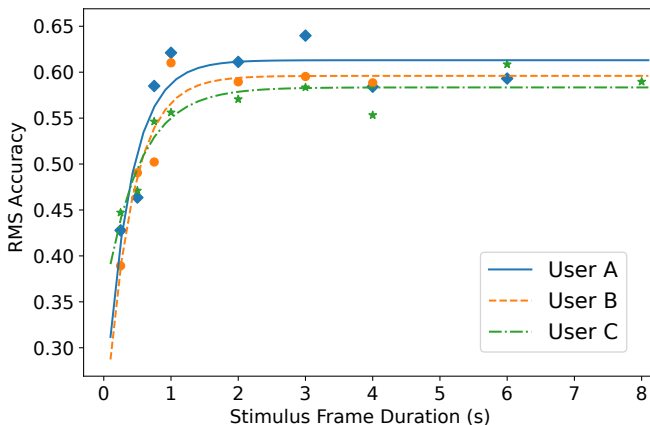

Fig. 17: RMS accuracy graph for the star using Low-passed LIA method at different stimulus frame duration for each user.

Table III highlights the best accuracy measurements for each user and method. A relatively long time duration produces better accuracy but not always. The table shows that the TNR has a stronger impact on RMS accuracy. The TPR for all users and methods is quite high with an average of about 0.989. The TNR, however, has a higher variance between methods. Integrated LIA has the best TNR values, which translate to having the best RMS accuracy, but it also has the highest time duration requirements to achieve that accuracy. Multi-peak FFT lags behind the other two LIA methods as none of the users reached past 0.6 RMS accuracy.

Some trade-off may exist between the user's attentiveness and visual field reconstruction accuracy as stimulus frame duration increase further. Participants have reported attention issues at higher durations. They have noted that faster stimulus

TABLE III: Best accuracy measurements obtained at a particular stimulus duration for each user and method.

		User A	User B	User C
Low-passed LIA	Frame Duration [s]	3.0	1.0	6.0
	RMS Accuracy	0.640	0.610	0.609
	TPR	0.994	1.000	1.000
	TNR	0.696	0.651	0.645
Integrated LIA	Frame Duration [s]	3.0	4.0	6.0
	RMS Accuracy	0.680	0.670	0.675
	TPR	0.982	0.988	1.000
	TNR	0.750	0.787	0.767
Multi-peak FFT	Frame Duration [s]	2.0	1.0	6.0
	RMS Accuracy	0.567	0.582	0.556
	TPR	0.970	0.982	0.982
	TNR	0.582	0.614	0.528

frame duration allows users to see new images more often, while longer duration forces participants to look at the same image for longer. However, we did not explore this issue in detail since doubling the stimulus frame duration will double the experiment time requirements. As an example, to reconstruct the 20 × 26 px star at 8 s requires 69 min of experiment time but a 4 s requires only about 34 min. For a low-resolution image like the star, this experiment is feasible but becomes exponentially harder as the image scale increases.

IV. CONCLUSION

We successfully designed an experimental process that maximizes signal quality during data collection. We also implemented improvements and algorithms which can better extract the source signal from EEG and SSVEP response data. Image reconstruction was completed with a low-cost, wearable, 5-channel EEG device, using only a single channel from the occipital lobe electrode. LIA algorithm implementations perform better than their FFT counterpart. In general, Low-passed LIA provided the best image reconstruction quality. With the new process and algorithms, multiple participants were able to take high definition "photographs" of multi-featured and complex objects such as the human face. We hope to further improve this technique by applying it in the real-world and using it to gain insight into human vision.

V. FUTURE WORK

A. Cognitive Studies and Visual Acuity

The results of this work may help us investigate links between the visual acuity of the human eye and the visual cortex. Cognitive-visual exercises, such as the one presented here, are useful to understand our capacity to sense our environment, extending to applications in mental health and personalized health care.

B. Image Reconstruction Enhancements

This paper presents signal processing techniques to reconstruct images from the human eye. Potentially, we can further enhance the reconstructed image by taking into account the sources of speckle noise. By combining the LIA and FFT processing with pattern analysis and machine learning techniques, we may improve upon the resulting image quality.

C. Multicolored Image Capture

Our techniques currently use black and white images. Studies have shown that red, green, and blue solid colors can be mapped based on SSVEP response amplitudes [39]. A possible next step is to map how our brain interprets solid colors that have red, green, and blue mixtures before producing a multicolored image capture. In this way, capturing real-life full-color representations from the human eye can be realized.

VI. ACKNOWLEDGEMENTS

REFERENCES

[1] W. R. Hendee and P. N. Wells, *The perception of visual information*. Springer Science & Business Media, 1997.

[2] A. M. Norcia, L. G. Appelbaum *et al.*, "The steady-state visual evoked potential in vision research: a review," *Journal of vision*, vol. 15, no. 6, pp. 4–4, 2015.

[3] S. Chen, H. Lee *et al.*, "A wireless body sensor network system for healthcare monitoring application," in *2007 IEEE Biomedical Circuits and Systems Conference*, 2007, pp. 243–246.

[4] S. Mann *et al.*, "Wearable, tetherless computer–mediated reality: WearCam as a wearable face–recognizer, and other applications for the disabled," in *Presentation at the American Association of Artificial Intelligence, 1996 Symposium. Retrieved January*, vol. 21, 1996, p. 2002.

[5] Y. Cai, J. Chen *et al.*, "Smart prosthesis system: Continuous automatic prosthesis fitting adjustment and real-time stress visualization," in *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 2018, pp. 1–4.

[6] J. Andreu-Perez, D. R. Leff *et al.*, "From wearable sensors to smart implants—toward pervasive and personalized healthcare," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 12, pp. 2750–2762, 2015.

[7] J. Ponmozhi, C. Frias *et al.*, "Smart sensors/actuators for biomedical applications," *Measurement*, vol. 45, no. 7, pp. 1675–1688, 2012.

[8] S. Mann, "Wearable computing as means for personal empowerment," in *Proc. 3rd Int. Conf. on Wearable Computing (ICWC)*, 1998, pp. 51–59.

[9] ——, "Wearable computing: A first step toward personal imaging," *Computer*, vol. 30, no. 2, pp. 25–32, 1997.

[10] R. Swaminathan and S. Prasad, "Brain computer interface used in health care technologies," in *Next Generation DNA Led Technologies*. Springer, 2016, pp. 49–58.

[11] A. Teles, M. Cagy *et al.*, "Using brain-computer interface and internet of things to improve healthcare for wheelchair users," in *UBICOMM 2017: The Eleventh International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, 2017.

[12] J. R. Wolpaw, N. Birbaumer *et al.*, "Brain-computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[13] S. Mann, *Humanistic Intelligence as a Basis for Intelligent Image Processing*. Wiley-IEEE Press, 2002.

[14] ——, "Humanistic intelligence," *Proc. Ars Electronica*, pp. 8–13, 1997.

[15] ——, "Surveillance (oversight), sousveillance (undersight), and metaveillance (seeing sight itself)," *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2016.

[16] S. Mann, S. Feiner *et al.*, "Wearable computing, 3D aug* reality, photographic/videographic gesture sensing, and veillance," in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, 2015, pp. 497–500.

[17] R. Janzen and S. Mann, "Sensory flux from the eye: Biological sensing-of-sensing (veillametrics) for 3D augmented-reality environments," *2015 IEEE Games Entertainment Media Conference (GEM)*, 2015.

[18] V. Mihajlović, B. Grundlehner *et al.*, "Wearable, wireless EEG solutions in daily life applications: what are we missing?" *IEEE journal of biomedical and health informatics*, vol. 19, no. 1, pp. 6–21, 2014.

[19] G. Shen, K. Dwivedi *et al.*, "End-to-end deep image reconstruction from human brain activity," *Frontiers in Computational Neuroscience*, vol. 13, p. 21, 2019.

[20] A. J. Casson, "Opportunities and challenges for ultra low power signal processing in wearable healthcare," in *2015 23rd European Signal Processing Conference (EUSIPCO)*. IEEE, 2015, pp. 424–428.

[21] "black-and-white-tree-clipart #3151774," Clipart Library, 2019. [Online]. Available: http://clipart-library.com/clipart/black-and-white-tree-clipart_11.htm

[22] R. C. Panicker, S. Puthusserypady, and Y. Sun, "An asynchronous p300 BCI with SSVEP-based control state detection," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, p. 1781–1788, 2011.

[23] G. Bin, X. Gao *et al.*, "An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method," *Journal of Neural Engineering*, vol. 6, no. 4, p. 046002, 3 2009.

[24] Y. Wang, X. Gao *et al.*, "Brain-computer interfaces based on visual evoked potentials," *IEEE Engineering in Medicine and Biology Magazine*, vol. 27, no. 5, p. 64–71, 2008.

[25] Y. Li, G. Bin *et al.*, "Analysis of phase coding SSVEP based on canonical correlation analysis (CCA)," *2011 5th International IEEE/EMBS Conference on Neural Engineering*, 2011.

[26] C. Jia, X. Gao *et al.*, "Frequency and phase mixed coding in SSVEP-based brain–computer interface," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, p. 200–206, 2011.

[27] S. A. Hillyard, H. Hinrichs *et al.*, "Combining steady-state visual evoked potentials and fMRI to localize brain activity during selective attention," *Human Brain Mapping*, vol. 5, no. 4, p. 287–292, 1997.

[28] C. S. Herrmann, "Human EEG responses to 1-100 hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Experimental Brain Research*, vol. 137, no. 3-4, p. 346–353, 2 2001.

[29] M. A. Pastor, J. Artieda *et al.*, "Human cerebral activation during steady-state visual-evoked responses," *The Journal of Neuroscience*, vol. 23, no. 37, p. 11621–11627, 2003.

[30] S. Mann, D. Lam *et al.*, "The human eye as a camera," in *2019 IEEE International Conference on E-health Networking, Application Services (HealthCom)*, 11 2019, p. 1–8.

[31] S. M. T. Müller, T. F. Bastos-Filho, and M. Sarcinelli-Filho, "Monopolar and bipolar electrode settings for SSVEP-based brain-computer interface," *Journal of Medical and Biological Engineering*, vol. 35, no. 4, pp. 482–491, 2015.

[32] "Simple panda cliparts #2475244," Clipart Library, 2019. [Online]. Available: http://clipart-library.com/clipart/701339.htm

[33] C. Van Loan, *Computational frameworks for the fast Fourier transform*. Siam, 1992, vol. 10.

[34] T. M. Peters and J. H. T. Bates, "The discrete fourier transform and the fast fourier transform," *The Fourier Transform in Biomedical Engineering*, p. 175–194, 1998.

[35] S. Devore, A. Gauthier *et al.*, "Improving students understanding of lock-in amplifiers," *2013 Physics Education Research Conference Proceedings*, 1 2014.

[36] J. H. Scofield, "Frequency-domain description of a lock-in amplifier," *American journal of physics*, vol. 62, no. 2, pp. 129–133, 1994.

[37] G. R. Müller-Putz, E. Eder *et al.*, "Comparison of DFT and lock-in amplifier features and search for optimal electrode positions in SSVEP-based BCI," *Journal of Neuroscience Methods*, vol. 168, no. 1, pp. 174–181, Feb. 2008. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0165027007004736

[38] M. Nikolova, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1/2, p. 89–97, 2004.

[39] S. Rasheed and D. Marini, "Classification of EEG signals produced by RGB colour stimuli," *Journal of Biomedical Engineering and Medical Imaging*, vol. 2, no. 5, pp. 56–56, 2015.