Analysis of Retinal Implants On Object Recognition Using Phosphene Distortion Model in Simulated Prosthetic Vision

Chitsein Htun,[1,†] Justin Kasowski,[2] Michael Beyeler,[2,3]
[1]*North Hollywood High School, 5231 Colfax Ave, North Hollywood, CA 91601*
[2]*Department of Dynamical Neuroscience, University of California, Santa Barbara, CA 93106*
[3]*Department of Computer Science, University of California, Santa Barbara, CA 93106*
[†]*corresponding author: chtun@live.com*

Degenerative retinal diseases such as retinitis pigmentosa and macular degeneration cause irreversible vision loss in over 10 million people worldwide. One treatment for these diseases includes retinal prostheses, which electrically stimulate surviving cells to induce neuronal responses that are interpreted by the brain as visual percepts ('phosphenes'). However, recent studies show that instead of seeing focal spots of light, implant users observe distorted phosphenes. By simulating prosthetic vision with a model characterizing these distortions, we analyze the efficacy of prosthetic devices (Argus II, PRIMA) in allowing users to recognize common objects. During this study, participants viewed and identified rotatable, freestanding objects through a filter simulating prosthetic vision. The results revealed an overall decrease in correct object identifications compared to results from models which do not consider phosphene distortion. In addition, there was a greater percentage of correct object identifications with PRIMA-based devices compared to Argus II devices with equivalent electrode counts. Overall, this research models and analyzes a more accurate portrayal of prosthetic vision and argues for the improvement of retinal implants to account for phosphene distortions.

Keywords: Visual Prosthesis, Phosphene Distortion, Simulated Prosthetic Vision, Virtual Reality

I. INTRODUCTION

Retinal prostheses are electronic implant devices that stimulate retinal cells to restore vision in individuals with degenerative retinal diseases, including retinitis pigmentosa and dry age-related macular degeneration (AMD) [1]. These diseases cause the loss of photoreceptors in outer retinal layers but leave inner retinal layer cells intact [2, 3]. By stimulating these surviving neuronal cells, retinal implant users observe visual percepts ('phosphenes'). In the past two decades, many advancements in retinal prosthesis technology have been made, but there are still several limitations to providing functional vision with retinal prostheses [1, 4].

Some of these limitations, such as low resolution, have been studied extensively [4-8], but other issues, namely phosphene distortions, have only recently been investigated. During electrode stimulation, the appearance of individual phosphenes is highly variable across users and electrodes within a user; users typically report seeing distorted and often elongated geometric shapes that fade quickly over time [9, 10]. As retinal implant technology continues to advance, higher electrode counts will increase resolution, but phosphene distortions will become a greater barrier to providing precise vision to users. To help understand these phosphene distortions, a recent study found that the position of electrodes along ganglion axon pathways in the retina accounts for these distortions [11]. The study concluded that when epiretinal stimulation occurs, electrodes activate neuronal cells in optic nerve axon fibers, which causes streaks in phosphenes ('phosphene tails').

In the past decade, the use of computer-based simulation has been helpful in understanding what implant users see [12] and developing software solutions to increase clarity [13]. However, the majority of past research done using simulated prosthetic vision has disregarded phosphene distortions, using a model referred to as the 'scoreboard model'[11]. The use of computer modelling to characterize distorted phosphenes requires the consideration of several parameters, including electrode position along the retina, activation sensitivity orthogonal to the axon (ρ, 'phosphene brightness and size'), and activation sensitivity along the axon (λ, 'distortion streak length')[11].

Through the simulation of prosthetic vision using the phosphene distortion models, we evaluate the parameters needed to achieve the ability to recognize objects while using prosthetic vision. In addition, we analyze and compare the different available retinal implant devices (Argus II from Second Sight Medical Products, PRIMA

from Pixium Vision) and different parameters in providing users with the ability to recognize objects.

II. METHODS

A. *Experiment Design*

The object recognition task was developed in Unity 2021.1.12f1, using C# and HLSL software to code the virtual scene and the prosthetic vision simulation filter.

Participants completed an object recognition task on a computer simulation program, run on personal computers at a resolution of 2560 × 1440 and a refresh rate of 60 Hz. The program captures the virtual scene information from a camera and processes it into images. To identify the objects, participants typed their answers into an input field. While typing, participants were shown autocomplete options from a list of over 984 common English nouns, including the objects presented in the task, to ensure proper spelling of participants' answers.

Prior to running the task, participants were seated in front of a computer monitor at a distance of 40-70 cm and given a keyboard and mouse. Participants completed two practice rounds of the simulation. During the first practice round, the prosthetic vision simulation filter was turned off, a 'chair' object was presented in the virtual scene, and participants were asked to identify the object presented. During the second practice round, the prosthetic vision simulation filter was turned on (settings: 31x19 electrodes, λ = 100, ρ = 50), a 'table' object was presented, and participants were asked to identify the object presented. After the two practice rounds, participants began the task.

Each participant was asked to identify objects 4 times for each of the 30 settings tested. The 30 settings were derived from all of the different combinations of the following electrode arrays, λ, and ρ values:

1. 10x6 electrode array (350 µm spacing between electrodes), 31x19 electrode array (350 µm spacing between electrodes), 50x50 electrode array (350 µm spacing between electrodes), 50x50 electrode array (250 µm spacing between electrodes), 100x100 (250 µm spacing between electrodes)
2. ρ = 100 µm, ρ = 300 µm
3. λ = 50 µm, λ = 1000 µm, λ = 5000 µm

The 10x6 electrode array models after the Argus II prosthetic device. The 31x19 electrode array models after the PRIMA prosthetic device as it has roughly the same number of electrodes as the PRIMA device, but it has the same rectangular configuration as the Argus II. The λ = 50 µm represents the 'scoreboard model'. Higher λ values (λ = 5000 µm) are used to represent higher levels of axonal stimulation which is typically found in epiretinal devices (Argus II). Lower λ values (λ = 50 µm) are used to represent lower levels of axonal stimulation which is typically found in subretinal devices (PRIMA) [14].

FIG. 1. The 22 objects presented to participants during the 'object recognition' task.

The 30 settings were presented randomly throughout the trial. Objects were presented one at a time and were randomly selected from a set of 22 objects shown in Figure 1. Participants were allowed to rotate the presented objects by using arrow keys, and could move the object closer and further from the camera using the scroll wheel. Participants were also allowed to change the angle of the camera by moving their mouse and could toggle on and off by using the escape key.

Throughout the experiment, the software recorded participants' accuracy and time to completion in seconds. Afterwards, the software adjusted the participants' time to completion based on a learning curve.

B. *Prosthetic Vision Simulation*

The simulation works through a combination of C# code processed by the CPU and fragment/compute shaders processed by the GPU. The axon map model is handled by a fragment shader, in which each pixel in our scene can be set to a specified RGBA (red, green, blue, alpha) value. This fragment shader requires five buffers to be set in GPU memory to process the scene. The first buffer stores the simulated electrode currents. These are calculated by another fragment shader run on the original image captured by the virtual camera. This shader, known as the preprocessing shader, works by comparing screen locations of the pixel currently being processed to the location on the screen of each simulated electrode (previously converted from retinal coordinates to screen coordinates). The image is converted to grayscale, and pixel locations are matched against the location of each electrode. This information is

stored in a registered buffer which is also passed into the axon model shader.

The second buffer required by the axon model contains each electrode's effect on each simulated axon segment. This is calculated by a compute shader and is based on the distance between simulated retinal locations of each axon segment and each electrode. The third, fourth and fifth buffers contain information for how much each axon segment affects each pixel, which is calculated by calling a python package, 'pulse2percept'. A contribution threshold is chosen by the user and a list of axon segments above this threshold (and their contribution) is generated for each pixel. The third and fourth buffers contain starting and stopping indexes for each individual pixel in the fifth buffer. This fifth buffer contains the amount of contribution for each axon segment that contributes to each individual pixel. These contributions could be sourced from any available phosphene model which contains information in a similar format (where each electrode activation is mapped onto a pixel). This allows for easy expandability in a highly optimized model as the majority of calculations are performed a single time at the start of the program. This also allows the user to add models for other types of devices that are not necessarily implanted on the retina. The results of these calculations can also be stored in a binary file and loaded almost instantaneously in subsequent runs with the same model.

The shader itself performs the minimal amount of calculations necessary to determine the brightness of each individual pixel. This brightness is determined by the maximum brightness contribution from all the axon segments that contribute to the pixel being processed. For each pixel, the shader loops through the buffer containing the contribution of axon segments that contribute to the pixel (starting and ending at the locations stored by the other two buffers). On each iteration of this loop, the axon segment's brightness contribution is calculated by looping through each electrode. On each iteration of this inner loop, the segment's brightness is incremented by the product of the information stored in the first two buffers (distance from axon segment to the current electrode and the current of this electrode) and the information in the third buffer (how much this axon segment contributes to the pixel being processed). The value calculated for this axon segment's brightness contribution is compared to the previous maximum brightness contribution and is stored if it is higher. After looping through each axon segment, the pixel's RGB values are set to the maximum brightness contribution. This is done for each individual pixel and an optional final shader is used to blur the resulting image.

	Electrode Array Configuration (*array dimensions (electrode spacing)*)					
	10x6 (350)	31x19 (350)	50x50 (350)	50x50 (250)	100x100 (250)	average accuracy
$\rho = 100$ $\lambda = 50$	1.0000	0.7500	1.0000	0.7500	1.0000	0.9000
$\rho = 300$ $\lambda = 50$	0.5000	0.7500	1.0000	1.0000	0.7500	0.8000
$\rho = 100$ $\lambda = 1000$	0.5000	0.5000	0.7500	1.0000	0.7500	0.7000
$\rho = 300$ $\lambda = 1000$	0.2500	1.0000	0.7500	0.7500	0.2500	0.6000
$\rho = 100$ $\lambda = 5000$	0.2500	0.7500	1.0000	0.7500	1.0000	0.7500
$\rho = 300$ $\lambda = 5000$	0.2500	0.2500	0.5000	0.5000	0.5000	0.4000
average accuracy	0.4583	0.6667	0.8000	0.7917	0.7083	

TABLE 1. Recorded accuracy of object recognition experiment for each setting. Accuracy for each setting and electrode array configuration were averaged.

III. RESULTS

The data collected shows an overall decrease in accuracy across all electrode arrays when distortions are applied to phosphenes. On average, the difference in accuracy between the scoreboard model and the phosphene distortion model was 25.00% (see Table 1). In addition, between simulated prosthetic vision with and without distortions, correct answers on average took 16.59% longer for participants viewing the object with distortions applied (see Table 2).

Higher λ values (23.75% average decrease in accuracy compared to settings with $\lambda = 50$ μm, see Table 1) played a more significant role in differences in accuracy compared with higher ρ values (18.33% average decrease in accuracy compared to settings with $\rho = 100$ μm, see Table 1). Higher λ values at low ρ values played a significant role in differences in time to completion (27.93% increase in time to completion compared to settings with $\rho = 100$ μm, see Table 2). However, higher ρ values played little to no role in differences in time to completion (8.08% decrease in time to completion compared to settings with $\rho = 100$ μm, see Table 2).

The PRIMA device simulation (31x19, $\lambda = 50$ μm) on average had a 50.00% greater accuracy and took 8.8602 fewer seconds to complete for correct answers than the Argus II device simulation (10x6, $\lambda = 5000$ μm). When electrode counts were equivalent (31x19), the PRIMA device simulation on average had a 25.00% greater accuracy (see Table 1) and took 3.6788 fewer seconds to

	Electrode Array Configuration (*array dimensions (electrode spacing)*)					
	10x6 (350)	31x19 (350)	50x50 (350)	50x50 (250)	100x100 (250)	average time to completion
$\rho = 100$ $\lambda = 50$	4.6588	6.7454	2.7648	8.7546	3.2409	5.2329
$\rho = 300$ $\lambda = 50$	11.0308	3.9492	3.7287	3.0893	12.9493	6.9495
$\rho = 100$ $\lambda = 1000$	7.7889	3.6775	2.2882	3.1201	4.1667	4.2083
$\rho = 300$ $\lambda = 1000$	12.1557	2.8923	4.6090	8.0713	1.8200	5.9097
$\rho = 100$ $\lambda = 5000$	19.2333	14.2671	3.2148	6.3868	2.8010	9.1806
$\rho = 300$ $\lambda = 5000$	9.1817	3.7850	1.5067	3.8462	2.9721	4.2583
average time to completion	10.6749	5.8860	3.0187	5.5447	4.6583	

TABLE 2. Recorded time to completion for correct answers in seconds of object recognition experiment for each setting. Time to completion for correct answers for each setting and electrode array configuration were averaged and adjusted based on a learning curve.

complete for correct answers (see Table 2) than the Argus II device simulation.

A limitation of the data is that it draws from a small participant pool. This was due to the high computational intensity of the experiment requiring the use of supercomputers at UCSB, and COVID-19 restrictions. More data collection would be useful to confirm results.

VI. DISCUSSION

In past studies, the majority of research has focused on object recognition in simulated prosthetic vision under low resolutions [4-8]. However, the results suggest that phosphene distortions also play a significant role in hindering the ability to recognize objects while using retinal prostheses. To achieve relative proficiency in object recognition (accuracy > 90%), it will require both high electrode counts (50x50 or higher), and low ρ and λ values (ρ = 100 μm, λ <= 50 μm).

The activation of ganglion axon pathways during retinal stimulation is the main cause of higher λ values [11]. To lower λ values, recent studies on visual acuity and perceived resolution have found that subretinal devices, the PRIMA device, in particular, may cause less stimulation of ganglion axon cells which lessens phosphene distortions,

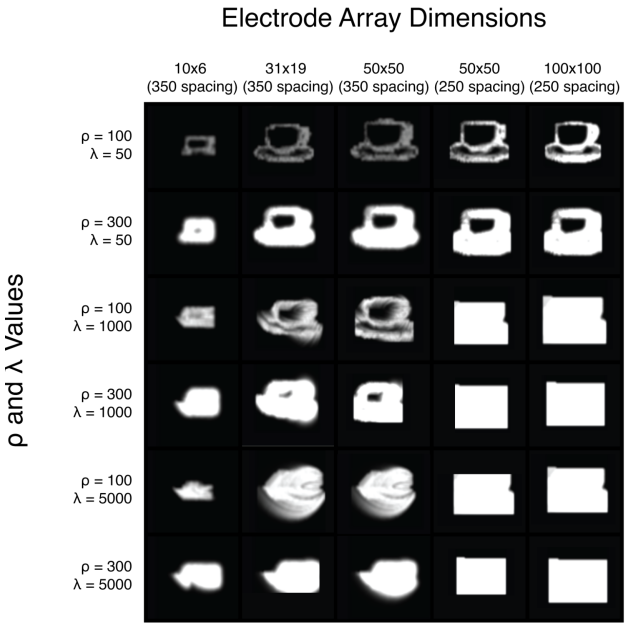

FIG. 3. Rendering of 'tea cup' object under prosthetic vision simulation at different electrode array dimensions and ρ and λ values at equivalent distance to the object.

when compared to epiretinal devices (Argus II) [14]. Additional studies will be required to find both hardware and software solutions to reduce λ values in subretinal devices further.

The electrode position relative to the stimulated retinal layers along with the strength of stimulation are the main factors of ρ values [11]. As electrode positions move further from stimulated retinal layers, ρ values subsequently increase, leading to enlargement of phosphenes and increased difficulty of recognizing objects [11]. In addition, as ρ values get larger, phosphene distortion streaks (λ) begin to affect the image less (see Figure 2). Our results suggest that a ρ value of 100 μm is an acceptable level for the tested electrode counts and electrode spacings. To achieve this ρ value, further research is needed on testing the effect of the distance from electrode to stimulated retinal layers and strength of stimulation on the ρ value.

The results confirm that differences in the ability to recognize objects between electrode counts fall off when comparing electrode counts of 50x50 or higher across all combinations of ρ and λ values [13].

Visual prostheses will be important devices in providing blind patients with artificial vision. Currently, there are several limitations to providing fully functional vision to users, including electrode counts and phosphene distortions. Overall, it will be essential to address both of these issues in order to improve the quality of artificial vision vastly.

ACKNOWLEDGEMENTS

The authors would like to thank Professor Michael Beyeler, Surya Jasper, and the UCSB Bionic Vision Lab for contributions to the simulated prosthetic vision model and software development and experiment development. The authors would also like to thank the UCSB Information Technology department for providing access to supercomputers to run the experiments on.

REFERENCES

[**1**] Ayton LN, Barnes N, Dagnelie G, et al. An update on retinal prostheses. Clin Neurophysiol. 2020;131(6):1383-1398.

[**2**] Ferrari S, Di Iorio E, Barbaro V, Ponzin D, Sorrentino FS, Parmeggiani F. Retinitis pigmentosa: genes and disease mechanisms. Curr Genomics. 2011;12(4):238-249.

[**3**] Mathenge W. Age-related macular degeneration. Community Eye Health. 2014;27(87):49-50.

[**4**] Eiber CD, Lovell NH, Suaning GJ. Attaining higher resolution visual prosthetics: a review of the factors and limitations. J Neural Eng. 2013 Feb;10(1):011002.

[**5**] Chang MH, Kim HS, Shin JH, Park KS. Facial identification in very low-resolution images simulating prosthetic vision. J Neural Eng. 2012 Aug;9(4):046012.

[**6**] Wang J, Li H, Fu W, Chen Y, Li L, Lyu Q, Han T, Chai X. Image Processing Strategies Based on a Visual Saliency Model for Object Recognition Under Simulated Prosthetic Vision. Artif Organs. 2016 Jan;40(1):94-100.

[**7**] Li S, Hu J, Chai X, Peng Y. Image recognition with a limited number of pixels for visual prostheses design. Artif Organs. 2012 Mar;36(3):266-74.

[**8**] Thorn JT, Migliorini E, Ghezzi D. Virtual reality simulation of epiretinal stimulation highlights the relevance of the visual angle in prosthetic vision. J Neural Eng. 2020 Nov 4;17(5):056019.

[**9**] Lu J, Mai G, Luo Y, Li M, Cao D, Wang X, Yan H, Sadda SR, Lu L. Appearance of Far Peripheral Retina in Normal Eyes by Ultra-widefield Fluorescein Angiography. Am J Ophthalmol. 2017 Jan;173:84-90.

[**10**] Yanai D, Weiland JD, Mahadevappa M, Greenberg RJ, Fine I, Humayun MS. Visual performance using a retinal prosthesis in three subjects with retinitis pigmentosa. Am J Ophthalmol. 2007 May;143(5):820-827.

[**11**] Beyeler M, Nanduri D, Weiland JD, Rokem A, Boynton GM, Fine I. A model of ganglion axon pathways accounts for percepts elicited by retinal implants. Sci Rep. 2019 Jun 24;9(1):9199.

[**12**] Chen SC, Suaning GJ, Morley JW, Lovell NH. Simulating prosthetic vision: I. Visual models of phosphenes. Vision Res. 2009 Jun;49(12):1493-506.

[**13**] Macé MJ, Guivarch V, Denis G, Jouffrais C. Simulated Prosthetic Vision: The Benefits of Computer-Based Object Recognition and Localization. Artif Organs. 2015 Jul;39(7):E102-13.

[**14**] Chenais, N.A.L., Airaghi Leccardi, M.J.I. & Ghezzi, D. Photovoltaic retinal prosthesis restores high-resolution responses to single-pixel stimulation in blind retinas. Commun Mater 2, 28 (2021).

APPENDIX

1. First Subsection

Figure 3a and 3b are 3D bar graph visualizations of Tables 1 and 2. Figure 3a shows an overall decrease in accuracy as distortion parameters increase. In addition, Figure 3a shows that higher electrode array dimensions show no significant improvement in accuracy, although more data collection may be necessary. Figure 3b shows an overall increase in time to completion as distortion parameters increase.

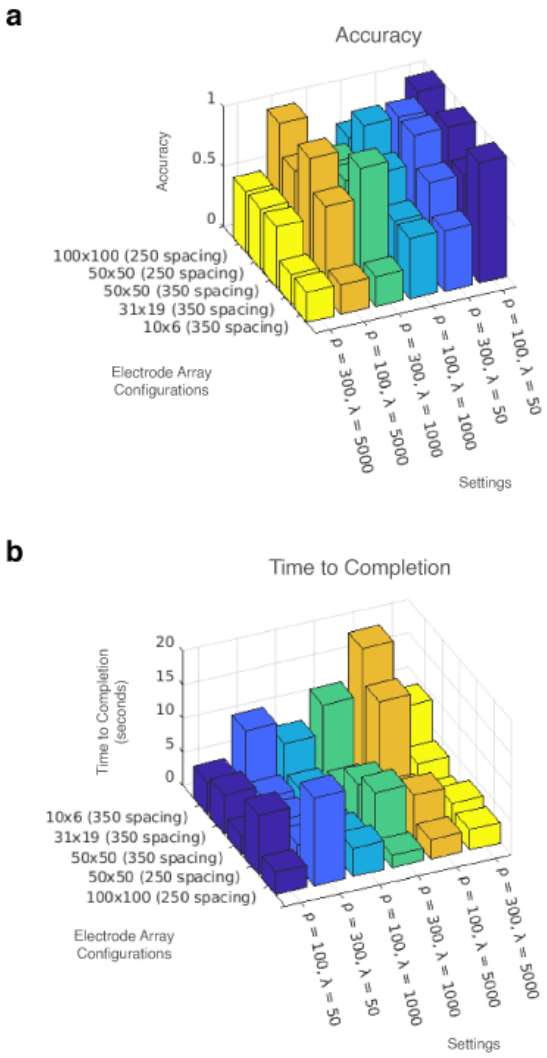

FIG. 3. (a) 3D bar graph of average accuracy as a function of electrode array configurations and ρ and λ settings. (b) 3D bar graph of average time to completion for correct answers in seconds as a function of electrode array configurations and ρ and λ settings.