Effects of Depth-Based Object Isolation in Simulated Retinal Prosthetic Vision

David Avraham * and Yitzhak Yitzhaky

Department of Electro-Optical Engineering, School of Electrical and Computer Engineering,
Ben-Gurion University of the Negev, Beer Sheva 84105, Israel; ytshak@bgu.ac.il
* Correspondence: davav@post.bgu.ac.il

Abstract: Visual retinal prostheses aim to restore vision for blind individuals who suffer from outer retinal degenerative diseases such as retinitis pigmentosa and age-related macular degeneration. Perception through retinal prostheses is very limited, but it can be improved by applying object isolation. We used an object isolation algorithm based on integral imaging to isolate objects of interest according to their depth from the camera and applied image processing manipulation to the isolated-object images. Subsequently, we applied a spatial prosthetic vision simulation that converted the isolated-object images to phosphene images. We compared the phosphene images for two types of input images, the original image (before applying object isolation), and the isolated-object image to illustrate the effects of object isolation on simulated prosthetic vision without and with multiple spatial variations of phosphenes, such as size and shape variations, spatial shifts, and dropout rate. The results show an improvement in the perceived shape, contrast, and dynamic range (number of gray levels) of objects in the phosphene image.

Keywords: retina; visual prosthesis; vision rehabilitation; perception; phosphene vision; integral imaging; object isolation

1. Introduction

Retinitis pigmentosa and geographic atrophy are outer retinal progressive degenerative diseases that cause a loss of photoreceptor cells. Both diseases currently have no cure [1,2] and together affect almost 0.7% of the global population [3,4]. One approach intended to help with blindness caused by outer retinal diseases is the use of a retinal prosthesis, which produces artificial electrical stimulation of the remaining healthy cells in the other layers of the retina.

During the late sixties, several researchers demonstrated that blind people can perceive electrically elicited light blobs, referred to as "phosphenes", by stimulating electrodes in contact with the occipital pole of the right cerebral hemisphere [5]. Since then, many visual prostheses have been developed. Retinal prostheses are a type of visual prostheses in which a microelectrode array is placed on the outer surface of the retina (subretinal prostheses), the inner surface of the retina (epiretinal prostheses), or the suprachoroidal layer (suprachoroidal prostheses), and remaining functioning cells of the retina are stimulated. Retinal prostheses aim to restore vision to blind individuals suffering from outer retinal progressive degenerative disorders, such as retinitis pigmentosa and age-related macular degeneration [6–11].

Spatial perception with current retinal prostheses is primarily limited by the low resolution (small number of electrodes). This has been addressed and demonstrated in most prior simulations [12–15]. Several other spatial variations of prosthetic vision have been reported. The first is phosphene size variation, which may be caused by variations in the stimulation amplitude, electrical conductivity, the efficiency of the electrodes [7,16–19], or different sizes of ganglion receptive fields [20]. The second is phosphene shape

variation. Phosphenes frequently appear round, but they may also appear as elongated shapes [8,17,21–23]; the elongated or elliptical shapes may be caused by unintended activation of axon bundles [20,24]. The phosphene shape may also be affected by the current amplitude of the stimulation [25]. The third factor influencing the perception is represented by dynamic random spatial shifts of the phosphenes. The spatial arrangement of the phosphenes in the subject's visual field was found to be correlated to the spatial arrangement of the electrodes [7,16,26–29]. However, this correspondence may be distorted by a mismatch between the electrode array and the underlying retinal ganglion cell (RGC) array [7,17].

Spatial variations have already been implemented in previous simulations. Dagnelie et al. [12] simulated variations in the sizes of phosphenes, the gap between phosphenes, varying number of electrodes (10×10, 16×16, and 25×25 electrodes), varying dynamic ranges (2, 4, and 8 gray levels), and different dropout rates (30, 50, and 70% of electrodes not eliciting a phosphene upon stimulation) and examined their impact on reading ability. Xia et al. [30] studied the simulated spatial shifts (phosphenes appear shifted with respect to the retinotopic position of the electrode array), variations in dynamic range (2, 4, 6, and 8), and different phosphene dropout rates (10, 20, 30, and 40%) on object recognition. Wu et al. [31] simulated spatial shifts, shape variations (50% of phosphenes reshaped into 12 different predefined elliptical shapes), and phosphene dropout (20%).

Epiretinal and suprachoroidal prosthetic devices use a video camera that captures the scene [6,10,11,32]. Image processing is then performed to fit this captured video to the prosthetic device. During this process, the 3D scene is converted into an appropriate 2D image. This conversion adds a distortion to the scene in addition to the spatial limitations of the prosthesis. Among these limitations are low resolution (small number of electrodes) [33], limited visual field [34], limited dynamic range (about 4–12 gray levels) [35], and cluttering. Cluttering is a situation in which an object of interest (OI) is confused or barely recognized because of the background or other objects in the original 3D scene.

Integral imaging (InI), originally proposed about a century ago [36], has become popular in the last two decades because of advances in digital imaging and optoelectronic technologies. In InI, 3D scene data are produced by capturing elemental images (EIs) optically using a pickup lens array and a detector array or by using multiple cameras arranged in a matrix form [37]. Integral imaging acquisition may be viewed as a multichannel system in which each channel generates an EI, which is a 2D image with its own perspective. These over-informative data permit the 3D visualization of objects. A 3D object isolation (or decluttering) algorithm using computational InI data has already been developed [38,39].

The limited number of retinal implant users (approximately 500 globally) [33] makes it difficult to conduct psychophysical studies with implant users and highlights the importance of simulating prosthetic vision. Prosthetic vision simulations are based on past psychophysical studies and vision research and should mimic the perception properties of retinal implants. These simulations demonstrate to normally sighted people how a blind person with a prosthetic device sees, and they enable the isolation of each spatial or temporal property and the investigation of its effects on prosthetic vision [40]. Simulations also help in examining how different image acquisition and image processing techniques affect prosthetic vision, which may also help in the specification of future developments of prosthetic devices and make them more practical for their users [35].

Because of the abovementioned spatial limitations and distortions in prosthetic vision, as well as several temporal distortions, such as persistence and perceptual fading of phosphenes [40,41], there is a strong demand for improving the quality of prosthetic vision. Here, we address this issue by applying an object isolation algorithm based on the depth information obtained from a computational InI, a 3D imaging technique that is currently not used in any of the available prosthetic devices. Subsequently, several image processing techniques were used to achieve additional improvement of the prosthetic images. This study presents results from multiple simulations.

The effects of object isolation in prosthetic vision were previously proposed and implemented in a binary representation [42]. However, this technique was not tested in a simulated prosthetic vision application. In this paper, we reveal the advantages of isolated-object input images compared with original scene input images (that contain multiple objects) in simulated prosthetic vision characterized by two phosphene models: symmetric phosphenes (i.e., without spatial variations) and asymmetric phosphenes (i.e., in the presence of multiple spatial variations). Our simulations are presented in both the original and reversed polarities (i.e., bright pixels are presented as dark and vice versa), and we discuss the advantages of these two presentations.

2. Methods

2.1. Scene Capture

In the experimental setup, a digital single-lens reflex (DSLR) camera (Nikon D500) was attached to a translator and took 100 EIs (10 rows of 10 EIs) by a shift and capture process. The horizontal and vertical translations between every two EIs were 5 mm. Each EI was taken at a resolution of 1856 × 2784 pixels, converted into grayscale, and then cropped and resized into a resolution of 500 × 1000 pixels (Figure 1) to reduce the algorithm running time. The captured scene is of a room that contains a laptop, a school bag, posters that were placed behind the school bag, and various less prominent objects. The laptop and the school bag were placed at distances of approximately 0.5 m and 1.2 m from the camera, respectively. The posters were placed behind the bag, approximately 2.3 m from the camera, to test the algorithm with a non-uniform background.

(**a**)　　　　　　　　　　　　　　　　　　　　　　　　　　　　　(**b**)
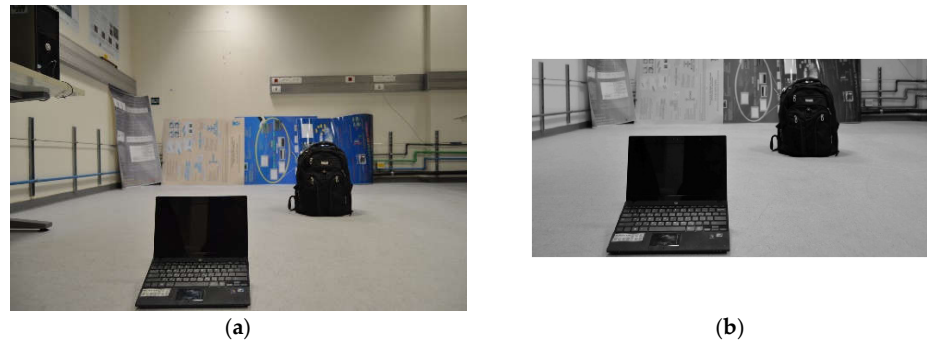
Figure 1. Preprocessing of the elemental images. (**a**) Original elemental image at a resolution of 1856 × 2784 and (**b**) a corresponding elemental image after preprocessing; the original image was converted into grayscale, cropped into a resolution of 1000 × 2000, and then resized into a resolution of 500 × 1000.

2.2. Object Isolation Algorithm

The benefits of object isolation on simulated prosthetic vision images were examined by adapting a method developed for depth-based 3D object isolation [38,39]. For this goal, computational integral imaging was performed using a matrix of EIs. The depth range available for the reconstruction of the depth planes was defined according to the depth of field of a single EI. A reconstructed image of the integral imaging system at z_0 depth is [37]:

$$f^{BP}(x, y, z_0) = \frac{1}{KL} \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} g_{kl}\left(x + \left(\frac{1}{M_0}\right) S_x k, y + \left(\frac{1}{M_0}\right) S_y l\right), \quad (1)$$

where g_{kl} is the $K \times L$ EI array; k and l are indices for the particular EI; M_0 is the magnification factor, which depends on the distance between the camera and the reconstructed plane z_0, S_x and S_y are the translations between two consecutive locations of the camera along x and y directions, respectively; and $f^{BP}(x, y, z_0)$ is a 2D

reconstructed image at a distance z_0 from the camera. Objects within the scene appear sharp when reconstructing the image at their depth, whereas objects far away from the chosen depth are blurred.

In the next phase, the 3D depth locations of objects in the 3D space were determined [39]. According to the assumption that the focused regions are obtained at higher frequencies [38], the higher local gradient values of the gradient image along the depth axis were determined:

$$f_{Grad}(x, y, z_0) = \nabla \cdot \left(f^{BP}(x, y, z_0) \right). \tag{2}$$

The threshold value was calculated according to the average gradient magnitude of the reconstructed images (AGMR) [39]:

$$AGMR(z_0) = \frac{1}{N_x \cdot N_y} \sum_y \sum_x |f_{Grad}(x, y, z_0)|, \tag{3}$$

where N_x and N_y are the numbers of pixels along x and y directions, respectively. Plotting the average gradient magnitude values against the depth locations on a graph produces local maxima in depths that include focused regions (i.e., locations of objects). Then, a spatial segmentation based on a sharpness criterion via an adaptive threshold was applied to the reconstructed images at the depths determined using the AGMR function [38,39].

The segmented planes obtained in this process are the depth-based isolated-object images intended to be the isolated-object inputs to the prosthesis. Another option for a prosthesis input can include all the isolated objects that may represent prominent objects in the observed scene. For this case, using the knowledge that computational integral imaging may observe regions hidden behind the front objects, and that the background in the isolated-object images is set to black in all the isolated-object images, the following procedure was carried out. The pixels that represented the front object (all the non-zero pixels) in the isolated-object image of the closest depth plane (to the camera) were identified. Next, the corresponding pixels were zeroed in the next-closest isolated-object images (i.e., the next-closest object's region hidden behind the one in the front). This process was repeated for all the isolated-object images of the farther depth planes until arriving at the isolated-object image of the farthest depth plane. Thus, in the case of spatial overlap between OIs, the algorithm uses the front object, similarly to normal vision. The following equation is used to produce an image that includes all the isolated objects:

$$I_{OIs} = \sum_{i=1}^{n} I_i \tag{4}$$

where I_{OIs} is the image that contains all the isolated objects of interest, n is the number of OIs in the scene (or the number of depth planes detected with prominent objects), and I_i is the isolated-object image at the *i*th depth plane.

2.3. Prosthetic Vision Simulation

A prosthetic vision simulation was built in MATLAB R2018a (Mathworks Inc., Natick, MA, USA). The simulation input was a grayscale image; the output was a simulated prosthetic vision image that includes discrete or overlapping phosphenes in grayscale and was presented on a desktop display.

Phosphenes usually appear round [8,16,17,21,29] and white or lightly colored [8,16,21]; thus, they were implemented accordingly. Phosphenes do not appear with sharp edges [12], and thus a two-dimensional elliptical Gaussian function was used to present a phosphene as follows:

$$f(x, y) = A exp\{-[a(x - x_0)^2 + 2b(x - x_0)(y - y_0) + c(y - y_0)^2]\}, \tag{5}$$

$$a = \frac{\cos^2\theta}{2\sigma_x^2} + \frac{\sin^2\theta}{2\sigma_y^2},$$

$$b = -\frac{\sin 2\theta}{4\sigma_x^2} + \frac{\sin 2\theta}{4\sigma_y^2},$$

$$c = \frac{\sin^2\theta}{2\sigma_x^2} + \frac{\cos^2\theta}{2\sigma_y^2},$$

where $f(x,y)$ is the pixel's value in the location (x,y); A represents the phosphene brightness at the center and is the mean value of a pixel block in the high-resolution camera's frame, which, in our case, was quantized into four normalized values (0, 1/3, 2/3, and 1); (x_0, y_0) is the phosphene's center position coordinates; σ_x and σ_y represent the spread of the phosphene along both axes; and the angle θ represents the clockwise rotation of the ellipsoid-shaped phosphene. If $\sigma_x = \sigma_y$, then b = 0, and the phosphene is round; otherwise, the phosphene is elliptical. Adding the spatial variations to Equation (5) gives:

$$
\begin{aligned}
f(x,y) = A\exp\Big\{-S\Big[a\big(x - (ix_0 + x_1)\big)^2 + 2b\big(x - (ix_0 + x_1)\big)\big(y - (jy_0 + y_1)\big) \\
+ c\big(y - (jy_0 + y_1)\big)^2\Big]\Big\},
\end{aligned}
\tag{6}
$$

where the indices i, j represent the coordinates of the phosphene's center point on the x and y axes, respectively; S represents the size variation; and x_1, y_1 represent the spatial shifts. If these variations are disabled in the simulation, then the values of S, x_1, y_1, and θ are equal to zero, and σ_x and σ_y are equal to 1. If the variations were enabled, A was equal to zero for 20% of the phosphenes (randomly), S was uniformly distributed between 0.5 (i.e., 50% smaller phosphene) and 1.5 (i.e., 50% larger phosphene), and x_1 and y_1 were uniformly distributed within a range from 0 to 4 pixels. σ_x was uniformly distributed between 1 and 3, σ_y was uniformly distributed between 1/3 and 1, and θ was uniformly distributed between zero (i.e., the phosphene is horizontal) and 180 (i.e., the phosphene is horizontal again after a clockwise rotation of 180 degrees). Table 1 summarizes the spatial parameters and describe them, and Table 2 provides the specific values for each parameter.

Table 1. The spatial parameters of the prosthetic vision simulation.

Spatial Parameter	Symbol	Description
Size variations	S	Random uniformly distributed variation of the phosphenes' sizes between 30% smaller than nominal and 30% larger than nominal.
Shape variations	$\sigma_x, \sigma_y, \theta$	Random change in the shape of 15% of the phosphenes into an ellipse with a random angle (0–180°) of rotation.
Spatial Shifts	x_1, y_1	Random uniformly distributed shifts of phosphenes by 0–4 pixels with respect to the retinotopic position of their corresponding electrodes in the array, in both the vertical and horizontal directions.
Dropout	A	Random elimination of 20% of the phosphenes.

Size variations and spatial shifts were applied to all phosphenes, shape variations were applied to only 15% because it is likely that only a few of the elicited phosphenes appear elliptical rather than round [7,16,17,21,22], and dropout (A) was applied to 20% of the phosphenes because it is unlikely that most of the electrodes are unable to evoke phosphenes. Figure 2 shows the different phosphene sizes and shapes.

Table 2. The values employed for the spatial parameters of the prosthetic vision simulation.

|		S	σ_x	σ_y	θ (deg)	x_1 (pxls)	y_1 (pxls)	A
Size variations	Disabled	1						
	Enabled	0.5–1.5						
Shape variations	Disabled		1	1	0			
	Enabled		1–3	1/3–1	0–180			
Spatial Shifts	Disabled					0	0	
	Enabled					0–4	0–4	
Dropout	Disabled							0, 1/3, 2/3, 1
	Enabled							0

(a)

(b)

(c)

(d)

(e)

(f)
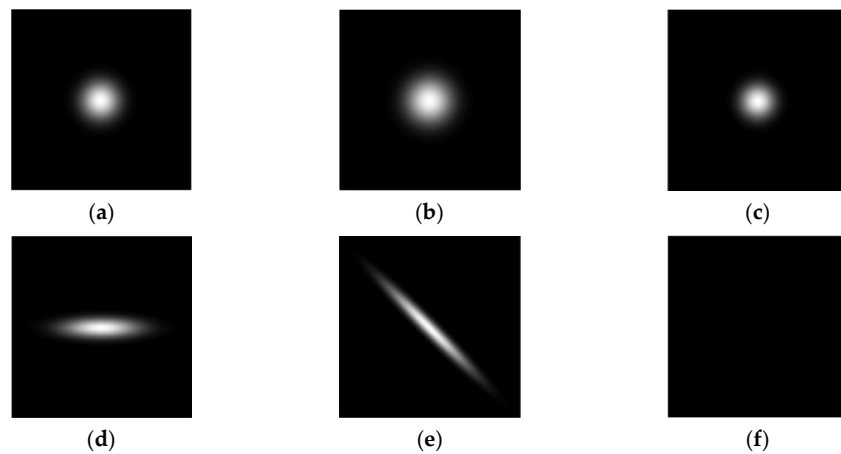
Figure 2. Examples of different sizes and shapes of phosphenes due to different simulation parameters. (**a**) $\sigma_x, \sigma_y = 1$, and $S = 1$. (**b**) $\sigma_x, \sigma_y = 1$, and $S = 1.3$. (**c**) $\sigma_x, \sigma_y = 1$, and $S = 0.7$. (**d**) $\sigma_x = 2, \sigma_y = 0.5, \theta = 0$, and $S = 1$. (**e**) $\sigma_x = 3, \sigma_y = 0.33, \theta = 45$ deg, and $S = 1$. (**f**) $A = 0$.

An electrode array that includes 15 × 30 electrodes in a rectangular structure was chosen to demonstrate the advantages of our technique on simulated prosthetic vision at a reasonable spatial resolution. Each stimulating electrode elicited one phosphene in the visual field. The phosphenes were rendered in a full 8-bit display, but the range of their maximum intensities was quantized into a 2-bit dynamic range (i.e., 4 different phosphene intensities).

The input image was divided into square pixel blocks so that the number of blocks matched the number and organization of the electrodes [40]. The original resolution of the input image (500 × 1000 pixels) was decreased to the low prosthetic vision resolution, which was set to 105 × 210 pixels to grossly fit the field of view of the Argus II prosthesis— 22° diagonally [34] on a head-mounted display (HMD) screen with a resolution of 960 × 1080 pixels and a field of view of 90° × 110° [43]. The average gray value of each pixel block in the image was calculated, and a phosphene was generated accordingly within the block region. The brightness of each phosphene was set based on the corresponding quantized value of the pixel block average, for which a higher average result in a brighter phosphene [40].

The phosphene images were rendered in two different ways: using discrete or overlapping phosphenes. The overlap rendering added uniformly distributed random size variations of phosphenes that resulted in overlaps among phosphenes. The simulation has a feature that enables switching between the original and reversed polarities of the input grayscale image, similarly to Argus II [44]. In the reversed polarity, each pixel has a new value equal to 255 minus the original value. A flow chart of the entire

process, including the object isolation algorithm and the prosthetic vision simulation, is presented in Figure 3.
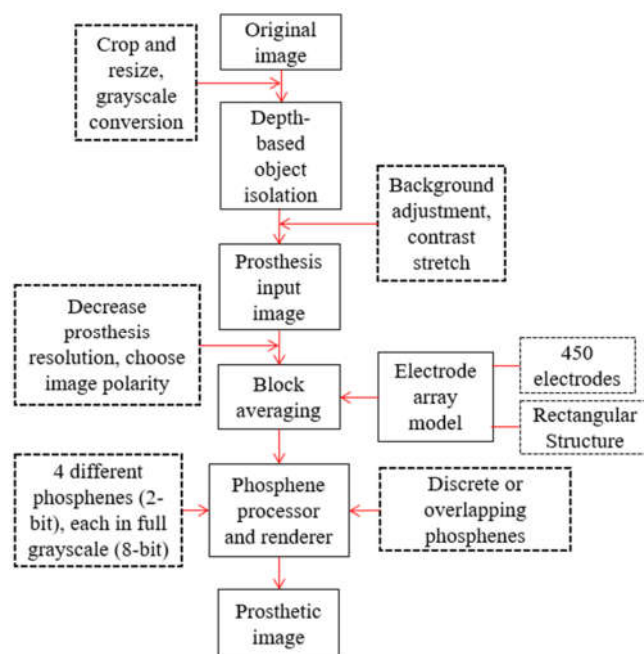

Figure 3. A flow chart of the simulation. The original image is cropped, resized, and converted to grayscale. Then, depth-based object isolation is applied. The background of the isolated-object image is chosen (black or white) based on the brightness of the object, and then, contrast stretch is applied. The prosthesis input image is rescaled to the prosthetic vision resolution and divided into pixel blocks according to the number of electrodes and the structure of the electrode array. The phosphene processor places a phosphene at the center of each block. The intensity of each phosphene is defined by the mean value of the pixels in the block. Four possible phosphene intensities are applied in the simulation, each presented in 8-bit grayscale. Either discrete or overlapping phosphenes are used to render the prosthetic image.

3. Results

3.1. Object Isolation

First, we determined the depths of the objects presented in Figure 1b. Figure 4 depicts the AGMR graph; the peaks in the graph represent the average gradient magnitude of the sharpest reconstructed images, which are at the depth locations of the significant objects in the 3D scene. The AGMR graph shows a peak at 190 mm and a significant change in the AGMR slope at 1190 mm. These values were correlated with the depths of the laptop and school bag, respectively. The reconstructed images at the depths identified in the AGMR graph are presented in Figure 5.
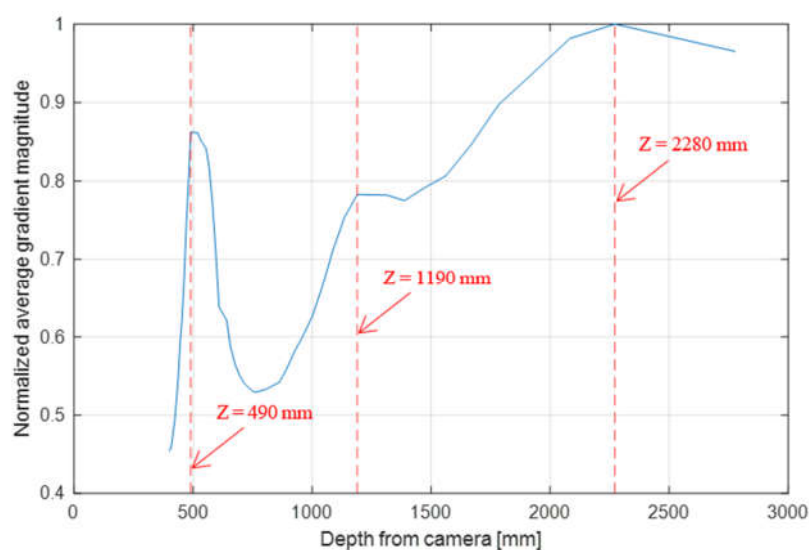
Figure 4. A graph of the average gradient magnitude of the reconstructed images of the scene presented in Figure 1. The horizontal axis is the distance from the camera in mm, ranging between 400 and 3000. The vertical axis is the absolute value of the gradient magnitude on the depth axis (z-axis). The graph has three peaks at 490, 1190, and 2280 mm, which correspond to the laptop, school bag, and posters behind the school bag, respectively.

(**a**)

(**b**)

(**c**)
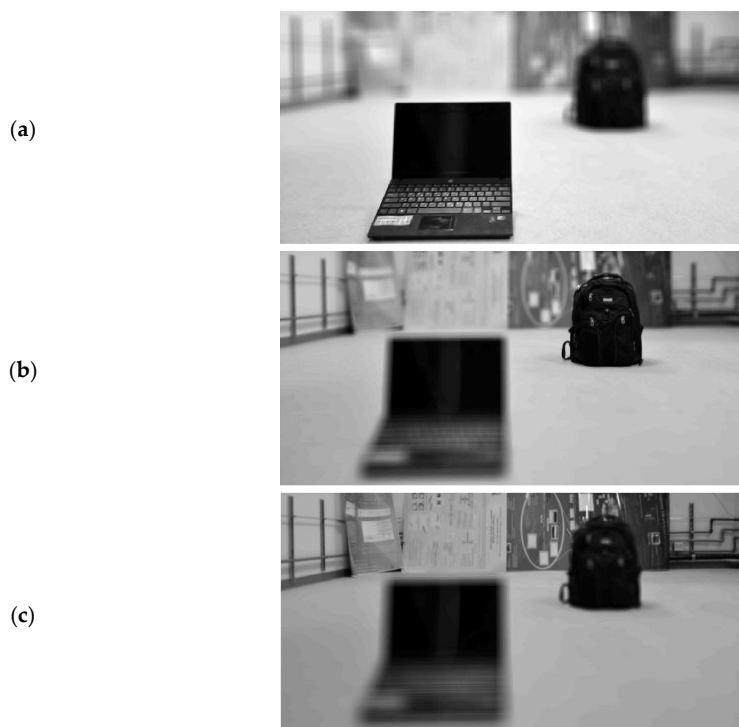
Figure 5. Reconstructed images according to the peak depths from the AGMR graph: (**a**) at z = 490 mm, (**b**) at z = 1190 mm, and (**c**) at 2280 mm.

3.2. Prosthetic Vision Views of the Isolated Objects

The rest of the object isolation process was performed [39]. The isolated-object images of the laptop and school bag were obtained, as presented in Figure 6. Direct prosthetic vision simulations of these images yielded a partial or no vision of the OIs because of the low dynamic range in prosthetic vision. After we applied contrast stretch to the isolated-

object images, the school bag became visible, but the laptop remained only partially visible. Then, we changed the black background to white before we applied contrast stretch. This made both objects visible under the low dynamic range and resolution conditions of the simulated prosthetic vision.
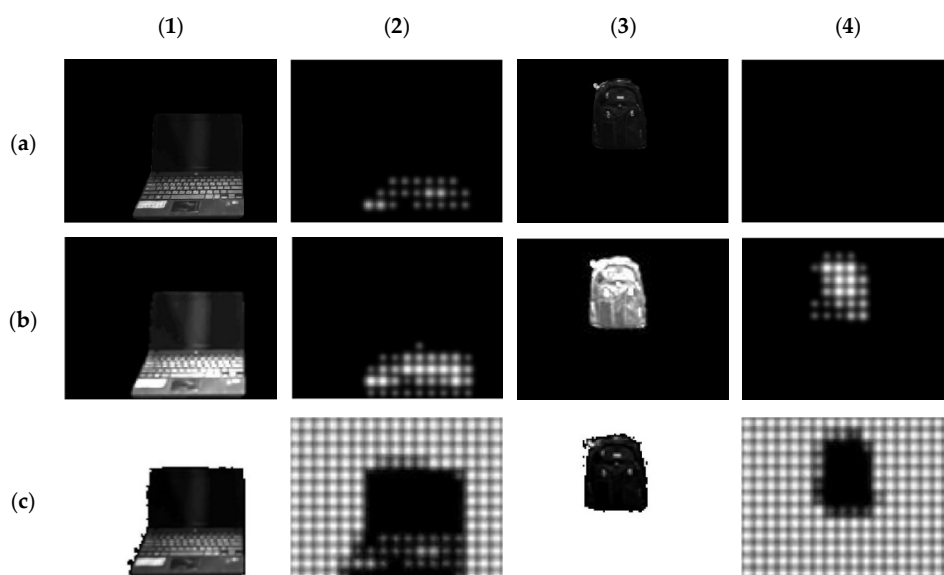

Figure 6. Isolated-object images based on a sharpness measure according to the peak depths from the AGMR graph and their corresponding simulated prosthetic image before and after image processing manipulations. Column 1: Reconstructed images of the laptop at z = 490 mm. Column 2: Simulated prosthetic images corresponding to the images from column 1 with a decreased resolution of 105 × 210 pixels. Column 3: Reconstructed image of the school bag at z = 1190 mm. Column 4: Simulated prosthetic images corresponding to the images from column 3 with a decreased resolution of 105 × 210 pixels. (**a**) Before applying contrast stretch; (**b**) after applying contrast stretch; and (**c**) after reversing the polarity (the background becomes white) and applying contrast stretch. Because of the decrease in the dynamic range from 256 gray levels to 4 gray levels, only a portion of the laptop remained visible, and the school bag entirely disappeared. Applying contrast stretch to the input images helps in perceiving the school bag, but the laptop remains only partially visible in prosthetic vision. With a white background input image, both the laptop and the school bag are visible in simulated prosthetic vision. All images are presented with an aspect ratio of 1:1.3 after partial cropping of the background.

3.3. Prosthetic Vision Views before and after Isolation

Figure 7 compares the simulated prosthetic vision images of symmetric phosphenes before and after the object isolation, background adjustment, and contrast stretch in both the original and reversed polarities. Figure 7a presents the original and simulated images of the scene from Figure 1b before applying any image processing manipulation to the input image. Figure 7b presents the original and simulated images of the isolated laptop from Figure 6(1)c. Figure 7c presents the original and simulated images of the isolated bag from Figure 6(4)c. Figure 7d presents the original and simulated images of the isolated OIs(according to Equation (4)).
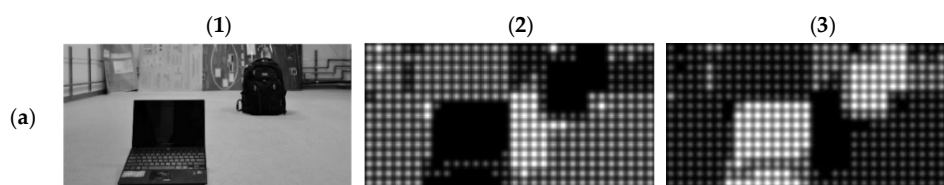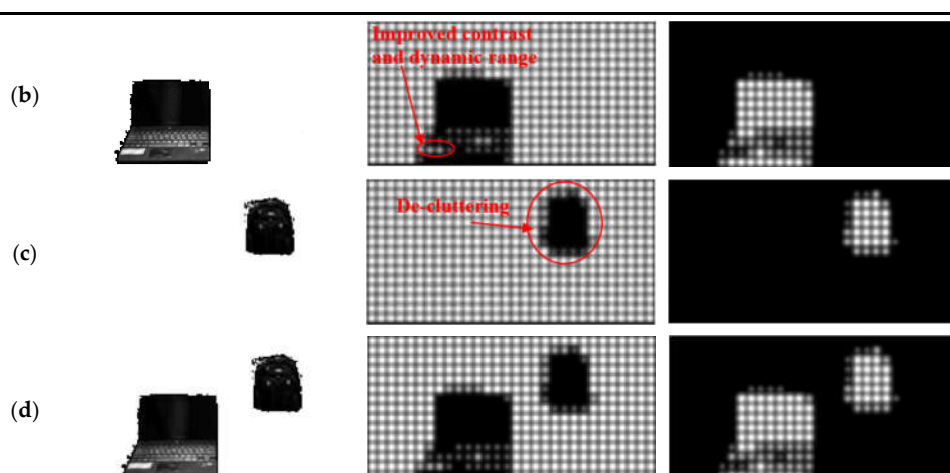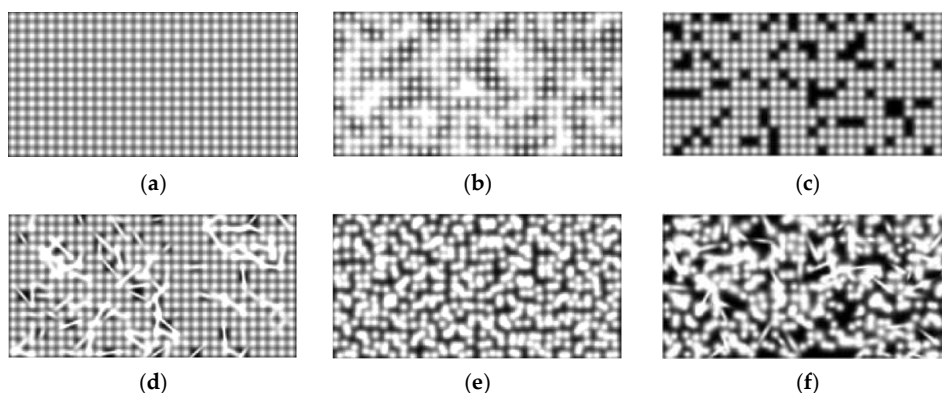
Figure 7. Simulated prosthetic vision images of symmetric phosphenes in two polarities. Simulated prosthetic images of (**a**) both OIs before the object isolation; (**b**) the laptop after the object isolation, background adjustment, and contrast stretch; (**c**) the school bag after the object isolation, background adjustment, and contrast stretch; and (**d**) both isolated OIs. Column 1: Input image. Column 2: Corresponding simulated prosthetic images in the original polarity. Column 3: Corresponding simulated prosthetic images in the reversed polarity. The object isolation and image processing manipulations resulted in an improved perception of the OIs compared with the results using the original input image in both image polarities in terms of their contrast, dynamic range, and perceived shape.

Figure 7b shows the contrast and dynamic range improvement that allows the perception of the white sticker on the laptop (in the red circle), as well as an improved perception of the keyboard and the screen. Figure 7c depicts the significant improvement in the perception of the school bag due to object isolation.

3.4. Prosthetic Vision Views before and after Isolation with Spatial Phosphene Variations

Figure 8 illustrates the effects of size variations, dropout, shape variations, and spatial shifts (non-uniform spacing) of phosphenes and a combination thereof in creating asymmetric phosphenes for a white input image. Figure 9 is the same as Figure 7 but with the addition of the spatial phosphene variations presented in Figure 8.

Figure 8. Visualization of various spatial phosphene variations. A simulated prosthetic image of a white input image with (**a**) no variations, (**b**) size variations of phosphenes, (**c**) dropout of phosphenes, (**d**) shape variations of phosphenes, (**e**) spatial shifts of phosphenes, and (**f**) all four spatial variations. The combination of all the spatial variations resulted in asymmetric phosphenes. The simulation is for an electrode array with 15 × 30 electrodes.

(1)   (2)   (3)
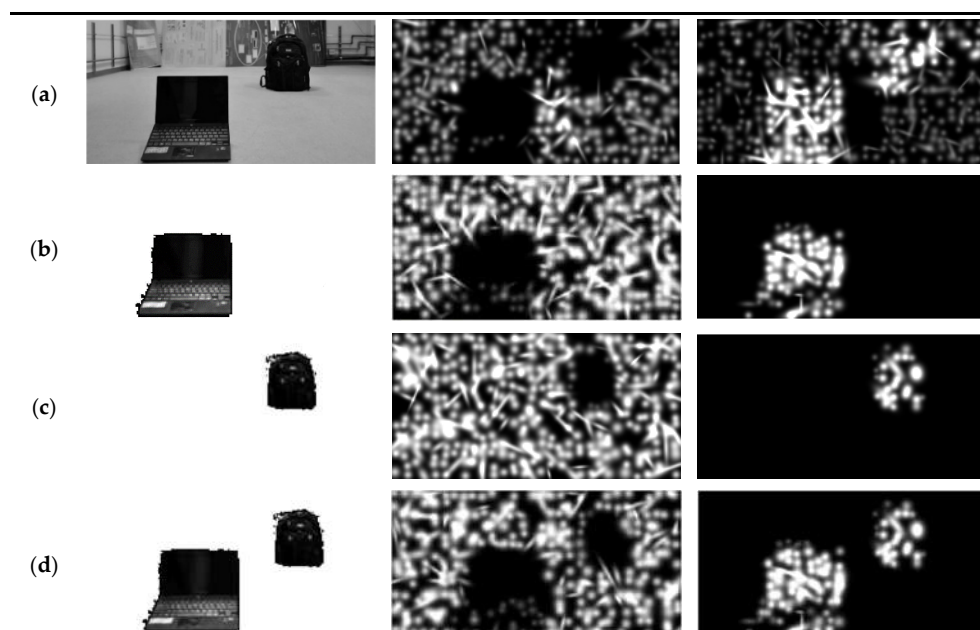
Figure 9. Simulated prosthetic vision images of asymmetric phosphenes in two polarities. Simulated prosthetic images including spatial variations of (**a**) both OIs before the object isolation; (**b**) the laptop after object isolation, background adjustment, and contrast stretch; (**c**) the school bag after object isolation, background adjustment, and contrast stretch; and (**d**) of both isolated OIs. Column 1: Input image. Column 2: Corresponding simulated prosthetic images in the original polarity. Column 3: Corresponding simulated prosthetic images in the reversed polarity. The simulated prosthetic images of the isolated-object images had improved contrast, dynamic range, and shape compared with the results using the original input image and with the addition of size and shape variations of phosphenes, spatial shifts, and dropout.

Figure 10 shows another example of the algorithm for a different scene. In this scene, the algorithm isolated the first two objects that were closest to the camera.
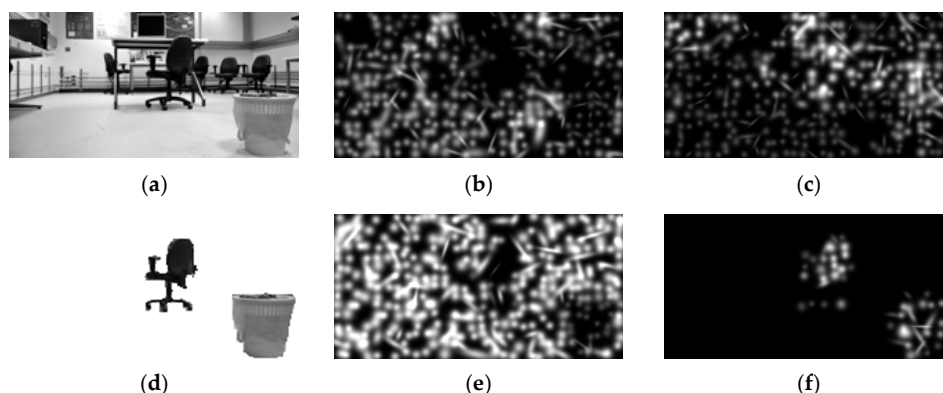
Figure 10. Simulated prosthetic vision images of asymmetric phosphenes before and after applying object isolation in a different scene. (**a**) Original image before the object isolation. (**b**,**c**) Corresponding simulated prosthetic images in the original and reversed polarities, respectively. (**d**) Input image after object isolation that shows the two OIs closest to the camera. (**e**,**f**) Corresponding simulated prosthetic images in the original and reversed polarities, respectively. Using the input image of isolated OIs instead of the original input image clearly improved the perception of the two closest objects in this scene.

4. Discussion and Conclusions

Despite the rapid progress in visual prosthetic devices, the spatial resolution remained low, with a visual acuity of 20/4275 for the Gen 2 suprachoroidal device [32], 20/1260 for the Argus II epiretinal prosthesis [28], and 20/546 for the ALPHA-IMS subretinal prosthesis [9]. Increasing the number of electrodes to improve the spatial resolution is challenging because of the limited power dissipation [45], crosstalk between electrodes [45,46], tissue heating [45], etc. [45,47]. Additionally, using sequential (asynchronous) stimulation, which reduces the temporal noise with many electrodes, limits the stimulation frequency [46]. It has also been found that the number of electrodes has a limited contribution to visual performance [48,49]. In addition to the low spatial resolution, numerous spatial and temporal effects can reduce the quality of prosthetic vision [7,18,19,22,40,41,50]. Using image processing approaches, such as the method proposed in this study, may be a desirable solution to improve the clarity of objects in prosthetic vision and reduce some of the distracting spatial and temporal effects.

We demonstrated the potential advantages of object isolation based on InI in simulated prosthetic phosphene vision. As part of the object isolation algorithm, a segmentation process based on a sharpness criterion was implemented to isolate the OIs in their depth planes. Our results agree with the conclusion that the depth-based object isolation process enables improved perception of the OI [42] by generating an isolated-object input, which avoids overlaps between the OI and the background or other objects.

A subject implanted with an epiretinal prosthesis has a low perceived dynamic range, which was reported to be as small as 2–4 gray levels [35], suggesting that the retinal neurons distinguish between a small number of distant stimulation levels within the safe charge density, which is $0.35\ mC/cm^2$ for platinum electrodes [51,52] and $3\ mC/cm^2$ for iridium oxide electrodes [53]. In our prosthetic vision simulation, the number of gray levels decreased from the standard 256 to 4. Therefore, enhancing the contrast and dynamic range of the OI in the isolated-object input image by applying contrast stretch allows the prosthesis to use four stimulation currents that are as far from each other as possible.

In the isolated-object image, the whole image except for the OI becomes uniform, usually black or white. When choosing between a black or a white background, the algorithm should consider the brightness of the OI. As shown in Figures 7b and 9b, applying contrast stretch to the isolated-object input image at this stage affects the simulated prosthetic image by increasing the OI's contrast (light gray phosphenes are being perceived in addition to dark gray phosphenes) and dynamic range (three levels instead of two in our example). This may enable brightness changes inside the OI, such as the sticker on the laptop, to be distinguished.

The proposed method may involve automatic isolation of predefined objects, which are of interest to implanted patients. This could be achieved by an object classification algorithm using machine learning. The applied object isolation technique may allow users to control the depth of their vision and thus increase their performance in multiple-object scenes. For instance, controlling the depth of vision could be achieved by a scroll wheel that would enable the users to see only objects within a predefined depth range; the objects outside this range would not be visible [42] (Figure 7b,c and Figure 9b,c). Alternatively, the input image of the prosthesis could contain objects at various depths, whereas other less prominent information associated with the background of the scene are removed (Figure 7d and Figure 9d).

Applying polarity reversal, in which the object is bright and the background is dark, has the advantage of involving fewer electrodes that stimulate bright phosphenes, thus decreasing the spread of distortions caused by spatial variations, such as phosphene shape [20], size [7,8], as well as non-uniform spatial shifts between phosphenes [7,53], and temporal variations of phosphenes, such as persistence and perceptual fading of phosphenes, in retinal prosthetic devices [40].

Our results suggest that object isolation may improve prosthetic vision regardless of the spatial variations, that is, in the simulated prosthetic vision of both symmetric and asymmetric phosphenes. The improvement in the simulated prosthetic vision images demonstrated in our study suggests that the technique presented in this work may improve prosthetic vision, and it should be considered in the development of future prosthetic devices.

A possible future direction of this study involves conducting experiments with normally sighted subjects to measure the performance improvement in different visual tasks, such as object detection, recognition, and identification. Another future direction is implementing the simulation in real time with an HMD. This would require a decrease in the running time by optimizing the algorithm. Such a real-time simulation will enable the testing of algorithm performance in other visual tasks, such as navigation and object scanning.

Author Contributions: "Conceptualization, D.A. and Y.Y.; methodology, D.A. and Y.Y.; software, D.A.; validation, D.A.; formal analysis, D.A.; investigation, D.A.; resources, D.A.; data curation, D.A.; writing—original draft preparation, D.A.; writing—review and editing, Y.Y.; visualization, D.A.; supervision, Y.Y.; project administration, Y.Y.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript."

Conflicts of Interest: The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this manuscript. The authors declare no conflicts of interest.

References

1. Wang, A.L.; Knight, D.K.; Vu, T.-T.T.; Mehta, M.C. Retinitis Pigmentosa: Review of Current Treatment. *Int. Ophthalmol. Clin.* **2019**, *59*, 263–280, doi:10.1097/iio.0000000000000256.
2. Holz, F.G.; Strauss, E.C.; Schmitz-Valckenberg, S.; Campagne, M.V.L. Geographic Atrophy. *Ophthalmology* **2014**, *121*, 1079–1091, doi:10.1016/j.ophtha.2013.11.023.
3. Narayan, D.S.; Wood, J.P.M.; Chidlow, G.; Casson, R.J. A review of the mechanisms of cone degeneration in retinitis pigmentosa. *Acta Ophthalmol.* **2016**, *94*, 748–754, doi:10.1111/aos.13141.
4. Bandello, F.; Silva, R. *AMD: Age-Related Macular Degeneration*; Théa: Loures, Lisbon, 2010.
5. Brindley, B.Y.G.S.; Lewin, W.S. The Sensations Produced by Electrical Simulation of the Visual Cortex. *J. Physiol.* **1968**, *196*, 479–493.
6. Hornig, R.; Zehnder, T.; Velikay-Parel, M.; Laube, T.; Feucht, M.; Richard, G. The IMI Retinal Implant System. In *Artificial Sight*; Springer: New York, NY, USA, 2007.
7. Humayun, M.S.; Weiland, J.; Fujii, G.Y.; Greenberg, R.; Williamson, R.; Little, J.; Mech, B.; Cimmarusti, V.; Van Boemel, G.; Dagnelie, G.; et al. Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vis. Res.* **2003**, *43*, 2573–2581, doi:10.1016/s0042-6989(03)00457-7.
8. Horsager, A.; Greenberg, R.J.; Fine, I. Spatiotemporal Interactions in Retinal Prosthesis Subjects. *Investig. Opthalmol. Vis. Sci.* **2010**, *51*, 1223–1233, doi:10.1167/iovs.09-3746.
9. Stingl, K.; Bartz-Schmidt, K.U.; Besch, D.; Braun, A.; Bruckmann, A.; Gekeler, F.; Greppmaier, U.; Hipp, S.; Hörtdörfer, G.; Kernstock, C.; et al. Artificial vision with wirelessly powered subretinal electronic implant alpha-IMS. *Proc. R. Soc. B Biol. Sci.* **2013**, *280*, 20130077, doi:10.1098/rspb.2013.0077.
10. Luo, Y.H.-L.; da Cruz, L. The Argus® II Retinal Prosthesis System. *Prog. Retin. Eye Res.* **2016**, *50*, 89–107, doi:10.1016/j.preteyeres.2015.09.003.
11. Roessler, G.; Laube, T.; Brockmann, C.; Kirschkamp, T.; Mazinani, B.; Goertz, M.; Koch, C.; Krisch, I.; Sellhaus, B.; Trieu, H.K.; et al. Implantation and Explantation of a Wireless Epiretinal Retina Implant Device: Observations during the EPIRET3 Prospective Clinical Trial. *Investig. Opthalmol. Vis. Sci.* **2009**, *50*, 3003–3008, doi:10.1167/iovs.08-2752.
12. Dagnelie, G.; Barnett, D.; Humayun, M.S.; Thompson, R.W. Paragraph text reading using a pixelized prosthetic vision simulator: Parameter dependence and task learning in free-viewing conditions. *Invest. Ophthalmol. Vis. Sci.* **2006**, *47*, 1241–1250.
13. Cha, K.; Horch, K.; Normann, R.A. Simulation of a phosphene-based visual field: Visual acuity in a pixelized vision system. *Ann. Biomed. Eng.* **1992**, *20*, 439–449, doi:10.1007/bf02368135.
14. Chai, X.; Yu, W.; Wang, J.; Zhao, Y.; Cai, C.; Ren, Q. Recognition of Pixelized Chinese Characters Using Simulated Prosthetic Vision. *Artif. Organs* **2007**, *31*, 175–182, doi:10.1111/j.1525-1594.2007.00362.x.

15. Guo, H.; Wang, Y.; Yang, Y.; Tong, S.; Zhu, Y.; Qiu, Y. Object Recognition Under Distorted Prosthetic Vision. *Artif. Organs* **2010**, *34*, 846–856, doi:10.1111/j.1525-1594.2009.00976.x.

16. Humayun, M.S.; de Juan, E., Jr.; Dagnelie, G.; Greenberg, R.J.; Propst, R.H.; Phillips, D.H. Visual Perception Elicited by Electrical Stimulation of Retina in Blind Humans. *Clin Sci.* **1996**, *114*, 40–46.

17. Rizzo, J.F.; Wyatt, J.; Loewenstein, J.; Kelly, S.; Shire, D. Perceptual Efficacy of Electrical Stimulation of Human Retina with a Microelectrode Array during Short-Term Surgical Trials. *Invest. Opthalmol. Vis. Sci.* **2003**, *44*, 5362–5369, doi:10.1167/iovs.02-0817.

18. Weiland, J.; Yanai, D.; Mahadevappa, M.; Williamson, R.; Mech, B.; Fujii, G.; Little, J.; Greenberg, R.; de Juan, E.; Humayun, M. Electrical stimulation of retina in blind humans. In Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439), Cancun, Mexico, 17–21 September 2004; Volume 3, pp. 2021–2022.

19. Nanduri, D.; Fine, I.; Horsager, A.; Boynton, G.M.; Humayun, M.S.; Greenberg, R.J.; Weiland, J. Frequency and Amplitude Modulation Have Different Effects on the Percepts Elicited by Retinal Stimulation. *Invest. Opthalmol. Vis. Sci.* **2012**, *53*, 205–214, doi:10.1167/iovs.11-8401.

20. Beyeler, M.; Nanduri, D.; Weiland, J.D.; Rokem, A.; Boynton, G.M.; Fine, I. A model of ganglion axon pathways accounts for percepts elicited by retinal implants. *Sci. Rep.* **2019**, *9*, 1–16, doi:10.1038/s41598-019-45416-4.

21. Mahadevappa, M.; Weiland, J.; Yanai, D.; Fine, I.; Greenberg, R.; Humayun, M. Perceptual thresholds and electrode impedance in three retinal prosthesis subjects. *IEEE Trans. Neural Syst. Rehab. Eng.* **2005**, *13*, 201–206, doi:10.1109/tnsre.2005.848687.

22. Yanai, D.; Weiland, J.D.; Mahadevappa, M.; Greenberg, R.J.; Fine, I.; Humayun, M.S. Visual Performance Using a Retinal Prosthesis in Three Subjects with Retinitis Pigmentosa. *Am. J. Ophthalmol.* **2007**, *143*, 820–827, doi:10.1016/j.ajo.2007.01.027.

23. De Balthasar, C.; Patel, S.; Roy, A.; Freda, R.; Greenwald, S.; Horsager, A.; Mahadevappa, M.; Yanai, D.; McMahon, M.J.; Humayun, M.S.; et al. Factors Affecting Perceptual Thresholds in Epiretinal Prostheses. *Invest. Opthalmol. Vis. Sci.* **2008**, *49*, 2303–2314, doi:10.1167/iovs.07-0696.

24. Grosberg, L.E.; Ganesan, K.; Goetz, G.A.; Madugula, S.S.; Bhaskhar, N.; Fan, V.; Li, P.; Hottowy, P.; Dabrowski, W.; Sher, A.; et al. Activation of ganglion cells and axon bundles using epiretinal electrical stimulation. *J. Neurophysiol.* **2017**, *118*, 1457–1471, doi:10.1152/jn.00750.2016.

25. Nanduri, D.; Humayun, M.; Greenberg, R.; McMahon, M.; Weiland, J. Retinal prosthesis phosphene shape analysis. In Proceedings of the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 20–25 August 2008; Volume 2008, pp. 1785–1788.

26. Caspi, A.; Dorn, J.D.; McClure, K.H.; Humayun, M.S.; Greenberg, R.J.; McMahon, M.J. Feasibility Study of a Retinal Prosthesis. *Arch. Ophthalmol.* **2009**, *127*, 398–401, doi:10.1001/archophthalmol.2009.20.

27. Sinclair, N.C.; Shivdasani, M.; Perera, T.; Gillespie, L.N.; McDermott, H.J.; Ayton, L.; Blamey, P.; for the Bionic Vision Australia Consortium. The Appearance of Phosphenes Elicited Using a Suprachoroidal Retinal Prosthesis. *Invest. Opthalmol. Vis. Sci.* **2016**, *57*, 4948–4961, doi:10.1167/iovs.15-18991.

28. Humayun, M.S.; Dorn, J.D.; da Cruz, L.; Dagnelie, G.; Sahel, J.-A.; Stanga, P.E.; Cideciyan, A.V.; Duncan, J.L.; Eliott, D.; Filley, E.; et al. Interim Results from the International Trial of Second Sight's Visual Prosthesis. *Ophthalmology* **2012**, *119*, 779–788, doi:10.1016/j.ophtha.2011.09.028.

29. Humayun, M.S.; de Juan E, Jr.; Weiland, J.D.; Dagnelie, G.; Katona, S.; Greenberg, R.; Suzuki, S. Pattern electrical stimulation of the human retina. *Vis. Res.* **1999**, *39*, 2569–2576.

30. Xia, P.; Hu, J.; Peng, Y. Adaptation to Phosphene Parameters Based on Multi-Object Recognition Using Simulated Prosthetic Vision. *Artif. Organs* **2015**, *39*, 1038–1045, doi:10.1111/aor.12504.

31. Wu, H.; Wang, J.; Li, H.; Chai, X. Prosthetic vision simulating system and its application based on retinal prosthesis. In Proceedings of the 2014 International Conference on Information Science, Electronics and Electrical Engineering, Sapporo, Japan, 26–28 April 2014; Volume 1, pp. 425–429.

32. Ayton, L.N.; Blamey, P.; Guymer, R.; Luu, C.; Nayagam, D.; Sinclair, N.C.; Shivdasani, M.; Yeoh, J.; McCombe, M.F.; Briggs, R.J.; et al. First-in-Human Trial of a Novel Suprachoroidal Retinal Prosthesis. *PLoS ONE* **2014**, *9*, e115239, doi:10.1371/journal.pone.0115239.

33. Ayton, L.N.; Barnes, N.; Dagnelie, G.; Fujikado, T.; Goetz, G.; Hornig, R.; Jones, B.W.; Muqit, M.M.; Rathbun, D.L.; Stingl, K.; et al. An update on retinal prostheses. *Clin. Neurophysiol.* **2020**, *131*, 1383–1398, doi:10.1016/j.clinph.2019.11.029.

34. Stronks, C., Dagnelie, G. The functional performance of the Argus II retinal prosthesis. *Expert Rev. Med. Devices* **2014**, *11*, 23–30.

35. Chen, S.C.; Suaning, G.J.; Morley, J.W.; Lovell, N.H. Simulating prosthetic vision: I. Visual models of phosphenes. *Vis. Res.* **2009**, *49*, 1493–1506, doi:10.1016/j.visres.2009.02.003.

36. Lippmann, G. Épreuves Réversibles Donnant La Sensation Du Relief. *J. Phys. Théor. Appl.* **1908**, *7*, 821–825.

37. Tavakoli, B.; Javidi, B.; Watson, E. Three dimensional visualization by photon counting computational Integral Imaging. *Opt. Express* **2008**, *16*, 4426–4436, doi:10.1364/oe.16.004426.

38. Aloni, D.; Yitzhaky, Y. Detection of Object Existence from a Single Reconstructed Plane Obtained by Integral Imaging. *IEEE Photon. Technol. Lett.* **2014**, *26*, 726–728, doi:10.1109/LPT.2014.2304883.

39. Aloni, D.; Yitzhaky, Y. Automatic 3D object localization and isolation using computational integral imaging. *Appl. Opt.* **2015**, *54*, 6717–6724, doi:10.1364/ao.54.006717.

40. Avraham, D.; Jung, J.; Yitzhaky, Y.; Peli, E. Retinal prosthetic vision simulation: Temporal aspects. *J. Neural. Eng.* **2021**, *18*, 460d9, doi:10.1088/1741-2552/ac1b6c.

41. Fornos, A.P.; Sommerhalder, J.; Da Cruz, L.; Sahel, J.-A.; Mohand-Said, S.; Hafezi, F.; Pelizzone, M. Temporal Properties of Visual Perception on Electrical Stimulation of the Retina. *Invest. Opthalmol. Vis. Sci.* **2012**, *53*, 2720–2731, doi:10.1167/iovs.11-9344.

42. Jung, J.-H.; Aloni, D.; Yitzhaky, Y.; Peli, E. Active confocal imaging for visual prostheses. *Vis. Res.* **2015**, *111*, 182–196, doi:10.1016/j.visres.2014.10.023.

43. Farahani, N.; Post, R.; Duboy, J.; Ahmed, I.; Kolowitz, B.J.; Krinchai, T.; E.; Monaco, S.; Fine, J.L.; Hartman, D.J.; Pantanowitz, L. Exploring virtual reality technology and the Oculus Rift for the examination of digital pathology slides. *J. Pathol. Inform.* **2016**, *7*, 22, doi:10.4103/2153-3539.181766.

44. Second Sight. *Argus II Retinal Prosthesis System Device Fitting Manual*; Second Sight: Pontotoc, MS, USA, 2013.

45. Palanker, D.; Vankov, A.; Huie, P.; A.; Baccus, S. Design of a high-resolution optoelectronic retinal prosthesis. *J. Neural Eng.* **2005**, *2*, S105–S120, doi:10.1088/1741-2560/2/1/012.

46. Moleirinho, S.; Whalen, A.J.; I.; Fried, S.; Pezaris, J.S. The impact of synchronous versus asynchronous electrical stimulation in artificial vision. *J. Neural. Eng.* **2021**, *18*, 51001, doi:10.1088/1741-2552/abecf1.

47. Loudin, J.D.; Simanovskii, D.M.; VijayRaghavan, K.; Sramek, C.K.; Butterwick, A.F.; Huie, P.; McLean, G.Y.; Palanker, D.V. Optoelectronic retinal prosthesis: System design and performance. *J. Neural. Eng.* **2007**, *4*, S72–S84, doi:10.1088/1741-2560/4/1/s09.

48. Behrend, M.R.; Ahuja, A.K.; Humayun, M.S.; Chow, R.H.; Weiland, J.D. Resolution of the Epiretinal Prosthesis is not Limited by Electrode Size. *IEEE Trans. Neural Syst. Rehab. Eng.* **2011**, *19*, 436–442, doi:10.1109/tnsre.2011.2140132.

49. Han, N.; Srivastava, S.; Xu, A.; Klein, D.; Beyeler, M. Deep Learning—Based Scene Simplification for Bionic Vision. *arXiv* **2021**, arXiv:2102.00297.

50. Weiland, J.D.; Humayun, M.S. Retinal prosthesis. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 1412–1424.

51. Brummer, S.B.; Turner, M.J. Electrical Stimulation with Pt Electrodes: II-Estimation of Maximum Surface Redox (Theoretical Non-Gassing) Limits. *IEEE Trans. Biomed. Eng.* **1977**, *24*, 440–443, doi:10.1109/TBME.1977.326179.

52. Cogan, S.F.; Troyk, P.R.; Ehrlich, J.; Plante, T.D. In Vitro Comparison of the Charge-Injection Limits of Activated Iridium Oxide (AIROF) and Platinum-Iridium Microelectrodes Stuart. *In Vitro* **2005**, *52*, 1612–1614.

53. Rizzo, J.F.; Wyatt, J.; Loewenstein, J.; Kelly, S.; Shire, D. Methods and Perceptual Thresholds for Short-Term Electrical Stimulation of Human Retina with Microelectrode Arrays. *Invest. Opthalmol. Vis. Sci.* **2003**, *44*, 5355–5361, doi:10.1167/iovs.02-0819.