ACS APPLIED OPTICAL MATERIALS

pubs.acs.org/acsaom

Forum Article

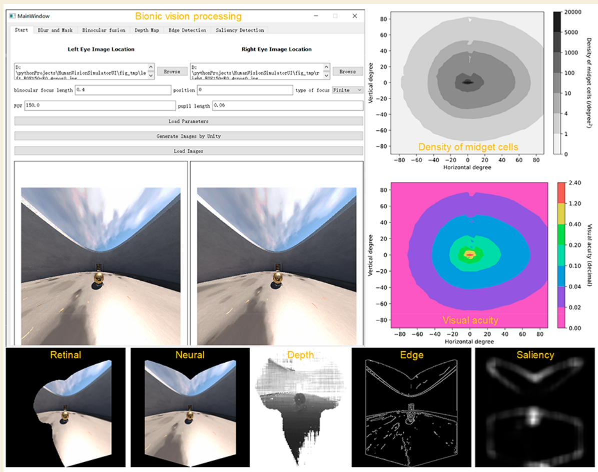
Bionic Vision Processing for Epiretinal Implant-Based Metaverse

Haiyang Hu, Chao Ping Chen,* Gang Li, Ziming Jin, Qiang Chu, Baoen Han, and Seak Pang Zou

Cite This: *ACS Appl. Opt. Mater.* 2024, 2, 1269−1276

Read Online

ABSTRACT: We present an epiretinal implant featuring bionic vision processing as a paradigm shift of metaverse. The main contribution of this work is to provide a methodology to better understand the human vision and to reproduce the stepwise images along the visual pathway. The epiretinal implant functions by stimulating the axons of ganglions to transmit the visual information to the brain. For the information on virtual environments to be correctly processed by the brain, our bionic vision processing is capable of transforming the digital images into neural images by factoring into the physiological pipelines of vision. The principles or algorithms of visual pathway, field of vision, visual acuity, foveated blurring, bilateral neural image fusion, depth perception, edge detection, and saliency detection are discussed. Our simulation results include the retinal images, neural images, depth map, edge map, and saliency map.



KEYWORDS: *retinal implant, ganglion, bionic vision, implantable device, virtual reality, metaverse*

1. INTRODUCTION

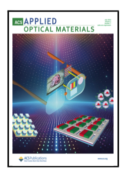
Throughout history, great philosophers have always been pondering over three fundamental questions. What are we? Where are we from? Where will we be? In the current mainstream view, we are human, or rather, *Homo sapiens*. We evolved from apes, who once lived in Africa. As to the third question, there are a lot of possibilities. In one scenario pictured by Prof. Yuval Noah Harari in his book *Homo Deus*,[1] *Homo sapiens* will be upgraded into cyborgs, a being with both organic and biomechatronic or bionic parts. This may sound like science fiction, but it is becoming a reality. As early as 2011, a retinal prosthesis called Argus II (Second Sight) began its clinical trials in Europe.[2] On May 25, 2023, Neuralink announced that it received approval from Food and Drug Administration to put chips into the human brain.[3] As the life expectancy continues to increase, it is inevitable that some parts of our body will fail and need replacements by artificial organs or implants, which could turn you into a cyborg without even realizing it. Furthermore, to cyborgs, the implants are not just body parts, but electronic devices such as the brain−computer interface. This would disrupt many things we know today. Metaverse, for example, can be accessed from implantable devices rather than wearable devices.[4−8] In the cyborg era, implantable devices may become more popular than wearable devices.

Inspired by the above prophecy, we hereby envision an implantable metaverse solution, which consists of two key components, i.e., epiretinal implant and bionic vision processing. In the subsequent sections, we will discuss the implant design, physiology of vision and principles of bionic vision, and substantiate our concept by reproducing the stepwise images along the visual pathway.

2. EPIRETINAL IMPLANT

Retinal implants or prostheses are a type of electronic implantable device for vision restoration.[9,10] Depending on where the implants are located, there are three major architectures: epiretinal, subretinal, and suprachoroidal.[11,12] Epiretinal implants are so far the most common, for it is easier and less risky to perform the surgery. In addition, by directly stimulating the ganglions, the epiretinal implants will require way fewer electrodes than the other two, given that the number of ganglions/cones/rods is 1.2/6/120 million, respectively.[13] Accordingly, the power consumption will also be dramatically reduced, which is a big plus for implantable devices.

Special Issue: Optical Materials and Applications for AR and VR Display Systems



Received: November 30, 2023
Revised: April 9, 2024
Accepted: April 11, 2024
Published: April 23, 2024

https://doi.org/10.1021/acsaom.3c00431
ACS Appl. Opt. Mater. 2024, 2, 1269−1276

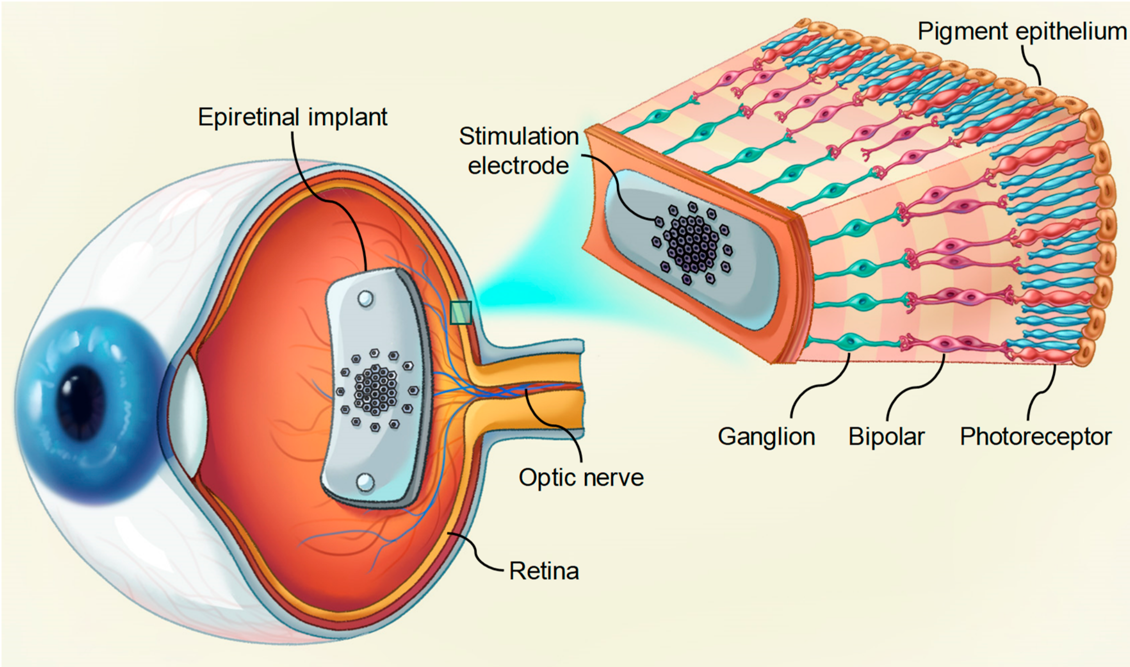


Figure 1. Schematic of the proposed epiretinal implant, where merely the ganglions, bipolar cells, photoreceptors and pigmented epithelium of retina are shown for the sake of simplicity. The implant is mounted on top of the layer of ganglions and has an array of electrodes that are able to stimulate the axons of ganglions, the bundle of which forms the optic nerve, to transmit the visual information to the brain.

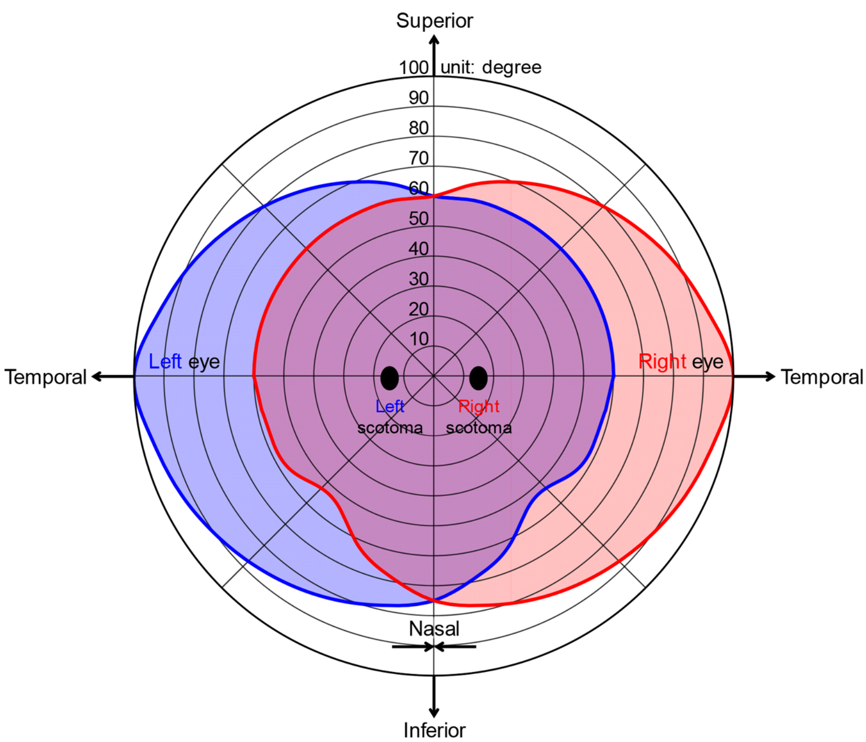


Figure 2. FOV in the polar coordinates. The blue/red area represents the FOV of left/right eye. The black area represents the blind spot. The data are derived from base values,[21,22] using linear interpolation. Yet, points near the nose are manually adjusted considering the obstruction of the nose.

Figure 1 is a schematic of the proposed epiretinal implant, where merely the ganglions, bipolar cells, photoreceptors and pigmented epithelium of retina are shown for the sake of simplicity. This implant is mounted on top of the layer of ganglions and has an array of electrodes that is able to stimulate the axons of ganglions, the bundle of which forms the optic nerve, to transmit the visual information to the brain. Ideally, the pattern or layout of stimulation electrodes shall match that of the ganglions. Surprisingly, the ganglions can be divided into various subtypes of cells.[14] To date, only three subtypes have been clearly identified. Subtype 1: midget cells or P-type ganglions. They account for about 80% of ganglions

https://doi.org/10.1021/acsaom.3c00431
ACS Appl. Opt. Mater. 2024, 2, 1269−1276

and play many important roles, including the color, spatial resolution, binocular parallax-induced depth perception,[13,14] etc. Subtype 2: parasol cells or M-type ganglions. In contrast to midget cells, parasol cells are larger in size and fewer in number (10% of ganglions). They are highly sensitive to luminance and hence associated with the flicker perception, motion perception, motion parallax-induced depth perception,[13,14] etc. Subtype 3: photosensitive ganglion cells. They constitute 1% of ganglions and are intrinsically photosensitive due to the melanopsin inside. They contribute to the circadian rhythm and pupillary light reflex.[13,14] Despite that all subtypes have their own physiological functions, we shall ignore the non-image-forming subtypes (e.g., parasol and photosensitive ganglion cells) and limit our interest to the midget cells for the bionic vision processing.

3. BIONIC VISION PROCESSING

3.1. Visual Pathway

To explain our bionic vision processing, we shall first go through the pipeline of the visual pathway, which is summarized as four steps. Step 1: light travels through the cornea, anterior and posterior chambers, pupil, lens, and vitreous chamber before reaching the retina.[15] Step 2: photoreceptors of the retina convert the light into the biological signals, i.e., the retinal images, via the visual phototransduction.[16] Step 3: retinal images are transmitted through the optic nerve (inside the optic tract), lateral geniculate nucleus, and optic radiation before reaching the primary visual cortex, during which the visual fields will be split at the chiasm and optic radiation. Step 4a: primary visual cortex percepts the color information carried by the midget cells.[17] Step 4b: primary visual cortex fuses two monocular images into a bilateral neural image. Step 4c: primary visual cortex detects the edge and saliency.[18] Step 4d: primary visual cortex percepts the binocular parallax-induced depth for the overlap of binocular fields.[19] For the brain to correctly process the visual information coming from the purely virtual environments, it is necessary to transform the digital or artificial images into the bionic images that are able to mimic the natural visual processing. Otherwise, visual discomfort or sickness may occur.[20]

3.2. Field of Vision

Each eye has a finite field of view (FOV). In each direction, monocular FOV has a maximum value of about 60°/75°/60°/100° with respect to the superior/inferior/nasal/temporal semimeridian, respectively.[21,22] As shown in Figure 2, the monocular FOV is 160° (horizontally) and 135° (vertically), while the binocular FOV is 120° (horizontally) and 135° (vertically), and the blind spot is located 15° (temporally) and 1.5° (inferiorly). By converting the polar coordinates to Cartesian coordinates, FOV in image coordinates can be plotted as in Figure 3, whose axes denote the relative length of an image. Besides, the area outside the blue/red area will be treated as a black image mask of left/right retinal image.

3.3. Visual Acuity

For a natural eye or vision, visual acuity shall involve both the optical and neural factors.[23] For our bionic vision, visual acuity shall depend on the neural factors only. In an ideal case, when no neural defects are found along the visual pathway, the visual acuity will solely hinge on the density of midget cells. With the raw data from a retinal topography calculator done by

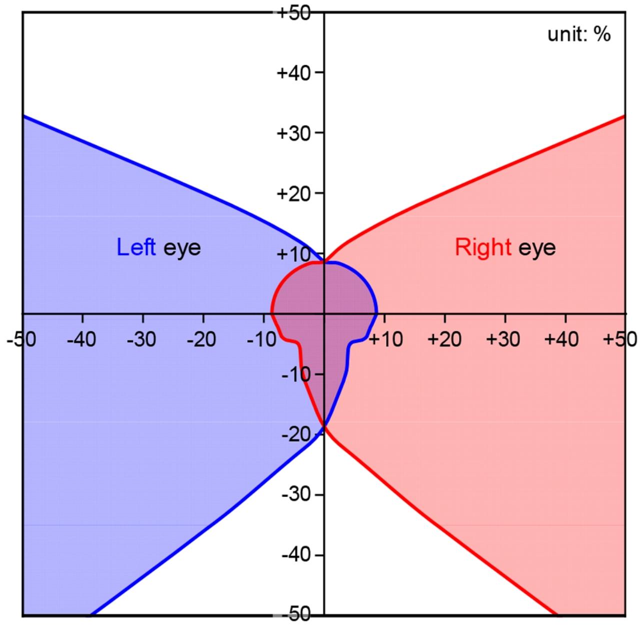


Figure 3. FOV in the image coordinates (i.e., Cartesian coordinates), whose axes denote the relative length of an image. The blue/red area represents the FOV of left/right eye. The area outside the blue/red area will be treated as a black image mask of left/right retinal image.

Watson,[24] the density of midget cells of the right eye could be plotted with respect to the visual field of up to ±90°, as shown in Figure 4, which can be further converted to the visual acuity, as shown in Figure 5, where the visual acuity is expressed as a decimal number and the best visual acuity is 2.2.

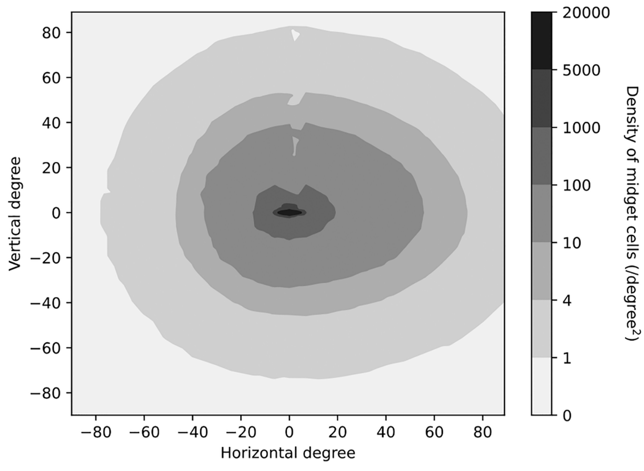


Figure 4. Density of midget cells of the right eye with respect to the visual field of up to ±90°. The raw data are extracted from a retinal topography calculator done by Watson.[24]

3.4. Foveated Blurring

Due to the uneven distribution of midget cells across the retina, there will be a foveated blurring effect to the retinal image. Since the image clarity or sharpness is proportional to the density of midget cells, the retinal image shall be blurred by applying different sizes of convolutional kernels at different retinal locations, which are measured by the offset angles from the fovea, as listed in Table 1. After weighing factors such as the complexity, time, and performance, a 15-level Gaussian

https://doi.org/10.1021/acsaom.3c00431
ACS Appl. Opt. Mater. 2024, 2, 1269−1276

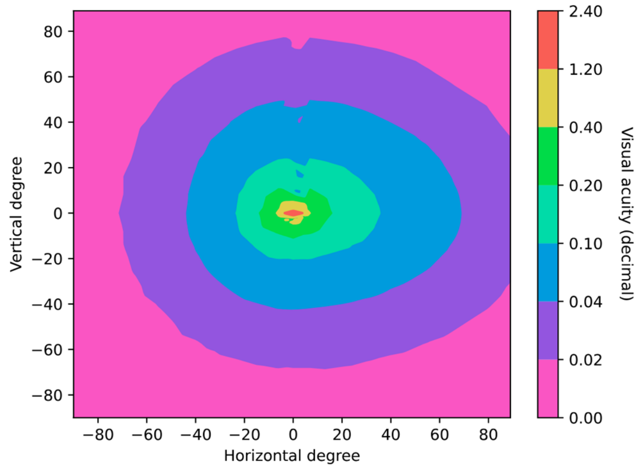


Figure 5. Visual acuity of the right eye with respect to the visual field of up to ±90°. The visual acuity is expressed as a decimal number and the best visual acuity is 2.2.

Table 1. Sizes of Convolutional Kernels for Blurring at Different Retinal Locations

level	kernel size	offset from fovea (degree)	offset from fovea (proportion)
1	1	25	0.0745
2	3	28	0.0840
3	5	31	0.0951
4	7	34	0.1062
5	9	37	0.1189
6	11	41	0.1363
7	13	43	0.1537
8	15	48	0.1759
9	17	52	0.2013
10	19	56	0.2345
11	21	61	0.2853
12	23	65	0.3391
13	25	70	0.4342
14	27	76	0.6355
15	29	81	1.0000

blurring is implemented. Compared to other blurring methods, where the density of midget cells is not considered,[25] the midget-cell-based foveation is arguably more accurate.

3.5. Bilateral Neural Image Fusion

When the binocular visual fields are to overlap, a conflict between the left and right brains will arise.[26] There are two ways to resolve this conflict: suppressing the image at one side of brain, or fusing left and right images. Either way, the phenomenon known as the eye dominance will come into play. As shown in Figure 6, it demonstrates the tendency of brains to prefer the visual information from the dominant eye. Statistically, the dominant eye of the majority is the right eye.[27]

3.6. Depth Perception

For depth perception, there exist two types of cues: binocular cues and monocular cues.[28] Unlike monocular cues, which rely upon the size, perspective line, occlusion, and motion parallax, etc., binocular cues primarily come from the binocular parallax or disparity.[29] As shown in Figure 7, the binocular parallax-based depth could be illustrated by an epipolar geometry and derived from

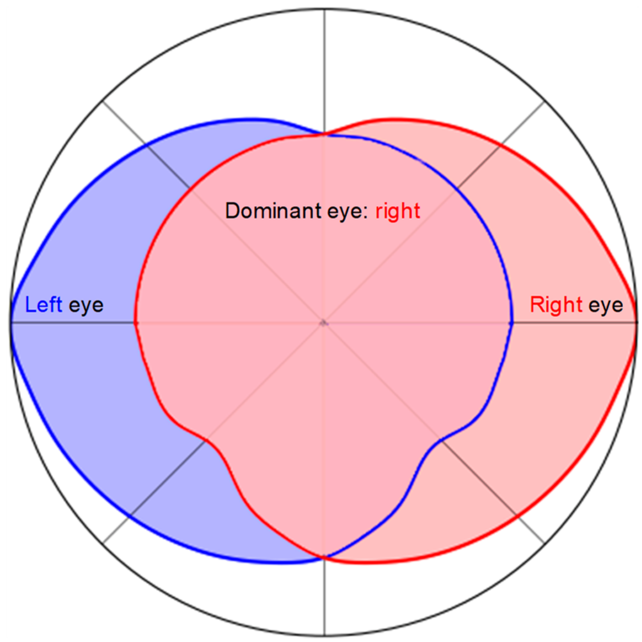


Figure 6. Right-eye dominance. When the binocular visual fields are to overlap, a conflict between the left and right brains will arise. Statistically, the dominant eye of the majority is the right eye.

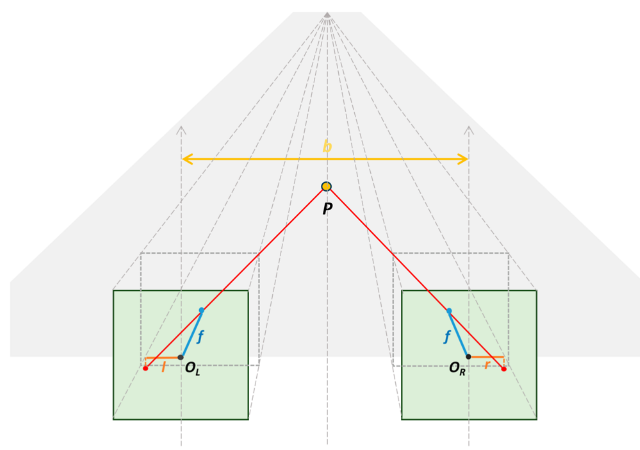


Figure 7. Geometry of epipolar correction for deriving the depth induced from the binocular parallax. P: point of interest. O_L/O_R: center point of the left/right retinal plane. b: interpupillary distance. f: focal length of the eye. l/r: bias from the center in the positive horizontal direction of the projections of P on the left/right retinal plane.

$$d = \frac{bf}{r - l} \tag{1}$$

where d is the depth of the point of interest, b the interpupillary distance, and f the focal length of the eye, and l/r the bias from the center in the positive horizontal direction of the projections of the point of interest in the left/right retinal plane. In the case of parallel eyes, the calculation of depth is straightforward. But when the viewing or gaze directions are unparallel, it is necessary to run the epipolar correction or image rectification prior to calculating the depth. This preprocessing is to project the images onto a common plane, thereby simplifying the process of identifying matching points in the two input images. Another important component

https://doi.org/10.1021/acsaom.3c00431
ACS Appl. Opt. Mater. 2024, 2, 1269−1276

of the depth perception is the parallax estimation or stereomatching, which aims to find the matched pairs of points in the left and right retinal images.

3.7. Edge Detection

Edge detection in the primary visual cortex is basically an image processing for finding the object boundaries, therefore allowing us to identify the shapes or recognize the objects. It is more sensitive to the relative difference in color and brightness.[30] To numerically model this function, the three most commonly used edge detection algorithms, i.e., Canny, Prewitt, and Laplacian of Gaussian detection, have been investigated.[31] A rule of thumb is that the Canny algorithm outperforms the other two due to its denoising, resistance to edge loss, and ability to preserve object details.[32,33] The process of Canny edge detection algorithm can be divided into five steps: Gaussian denoising, intensity gradient determination, non-maximum suppression, double thresholding, and edge tracking. Given that the edges can have multiple orientations, the detection is done with four filters of Sobel or other edge detection operators. Correspondingly, the orientation can be reduced to one of four angles (vertical, horizontal and two diagonals). Let Gx and Gy denote the first-order derivative of the image in the horizontal and vertical directions, respectively. The edge gradient G can be defined by the amplitude $|G|$

$$|G| = \sqrt{G_x^2 + G_y^2} \tag{2}$$

and the direction angle θ

$$\theta = \tan^{-1}(G_y/G_x) \tag{3}$$

Lastly, after several subsequent steps of enhancing and extracting the edges, the detection results can be obtained from the gradient information.

3.8. Saliency Detection

The primary visual cortex also generates the saliency maps to highlight the objects, to which our eyes are most likely to be focused.[34] To resemble this function, two static saliency detection algorithms are adopted.[35,36] The first is based on the visual saliency features, which leverages the integral of the original image to extract high-quality features.[35] The second is based on the spectral residual, which calculates the spectral residual of images in the spectral domain and accordingly plots the saliency map in the spatial domain. The saliency map can be calculated by

$$S(x) = g(x)F^{-1}(\exp[R(f) + P(f)])^2 \tag{4}$$

where $S(x)$ stands for the saliency map, $g(x)$ the Gaussian filter, F^{-1} the inverse Fourier transform, $R(f)$ the spectral residual, and $P(f)$ the phase spectrum of the image.[36]

4. EXPERIMENTAL DETAILS

4.1. Camera Settings

Unity is employed to create the three-dimensional (3D) scenes to emulate the physical worlds. Two physical cameras are deployed to imitate the binocular stereovision and to capture the left and right digital images as the inputs, as shown in Figure 8. To mimic the optics of eye,[37] the settings of the cameras are tweaked, as listed in Table 2, where the inter-camera distance (i.e., interpupillary distance) is 60 mm, focal length (i.e., diameter of the eyeball) is 25 mm, FOV is 150°, projected sensor size is 186 mm, and resolution (number of pixels) is 1860 × 1860.

Figure 8. (a) Left and (b) right digital images are captured by two physical cameras in Unity, where the intercamera distance is 60 mm, focal length is 25 mm, FOV is 150°, projected sensor size is 186 mm, and resolution is 1860 × 1860.

Table 2. Camera Settings To Capture Input Digital Images

parameter of camera	value
inter-camera distance	60 mm
focal length	25 mm
FOV	150°
projected sensor size	186 mm
resolution	1860 × 1860

4.2. Bionic Vision Processing

PyCharm (community edition 2021.1) serves as an integrated development environment for our bionic vision processing, which can transform the digital images into the bionic images, including the retinal image, neural image, depth map, edge map, and saliency map. Of various libraries, four shall be imported. OpenCV is to load an open-source computer vision library for image and video processing and analysis. Math is to access to a set of basic mathematical functions and constants. NumPy is to manipulate with multidimensional arrays and matrices. Matplotlib is to create data visualizations and graphs.

5. RESULTS AND DISCUSSION

5.1. Retinal Image

The retinal image, the image formed on the retina, can be obtained with four steps. First is to apply the 15-level foveated blurring to the input digital image. Second is to generate a black image mask for the given FOV, i.e., 150°. Third is to rectify both the foveated digital image and black image mask via the epipolar correction. Fourth is to merge the above two images into a retinal image, as shown in Figure 9, where the foveal or central region looks crisp, while the peripheral region gradually decreases in resolution.

5.2. Neural Image

The neural image, the image formed on the visual cortex, involves two types of image fusions: unilateral and bilateral. For the unilateral fusion, as shown in Figure 10, three steps are needed. First is to split the retinal image of each eye into left and right hemiretinal images. Since the blind spot is located near the bisector of the given FOV of 150°, left and right hemiretinal images are equally divided. The second is to move the left and right hemiretinal images to their respective visual pathways after the optic chiasm and to rectify them. The third is to merge the hemiretinal images into the unilateral neural image. The bilateral fusion, as shown in Figure 11, is to merge the left and right unilateral images via the aforementioned right eye dominance into the bilateral neural image, which is the product of cooperation between the left and right brains.

https://doi.org/10.1021/acsaom.3c00431
ACS Appl. Opt. Mater. 2024, 2, 1269−1276

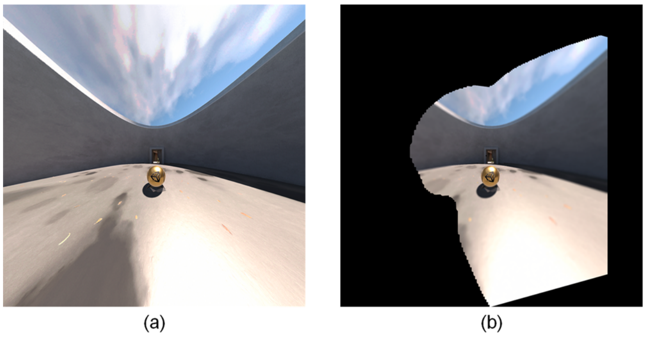


Figure 9. (a) Input digital image. (b) Retinal image, whose foveal or central region looks crisp, while the peripheral region gradually decreases in resolution.

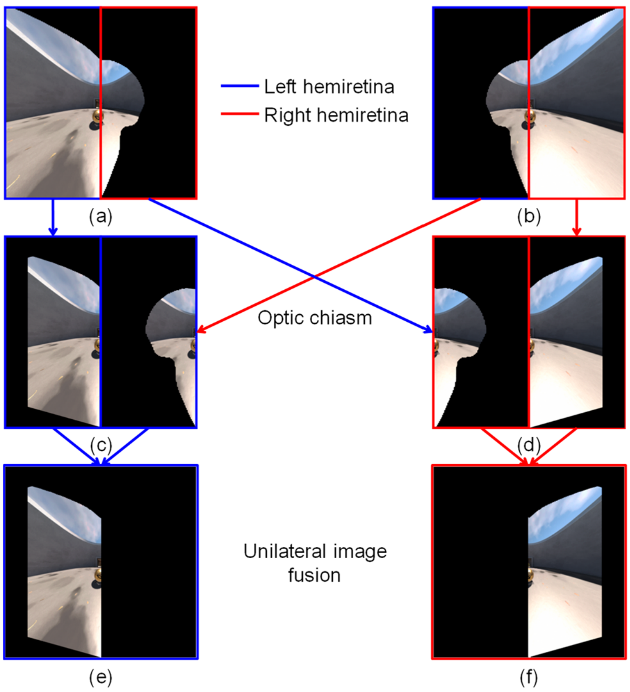


Figure 10. Unilateral neural image fusion. (a) Left retinal image. (b) Right retinal image. (c) Left hemiretinal images after optic chiasm. (d) Right hemiretinal images after optic chiasm. (e) Left unilateral neural image. (f) Right unilateral neural image.

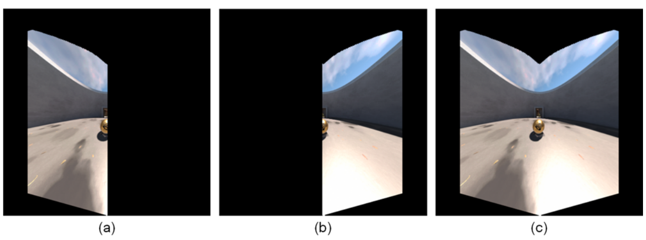


Figure 11. Bilateral neural image fusion. (a) Left unilateral neural image. (b) Right unilateral neural image. (c) Bilateral neural image.

5.3. Depth Map

To acquire the depth map, for which we shall limit the depth cue to the binocular parallax, four steps are implemented. Since the binocular parallax is color irrelevant, monochromatic images are used. The first is to rectify both left (Figure 12(a)) and right (Figure 12(b)) retinal images for easier point

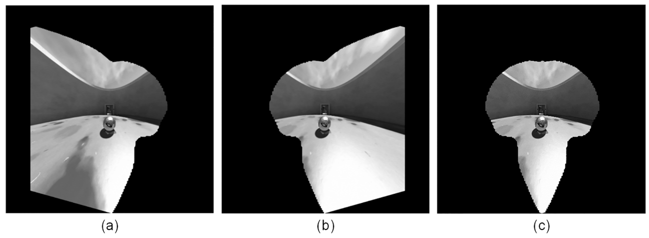


Figure 12. (a) Rectified left retinal image. (b) Rectified right retinal image. (c) Overlap of the rectified left and right retinal images.

matching. The second is to define the overlap region (Figure 12(c)) of the rectified left and right retinal images. The third is to match the points that correspond to the same point in the above overlap. The algorithm for point matching is called semiglobal block matching (SGBM), which is an intensity-based approach and generates a dense and smooth disparity map for 3D reconstruction.[38] The block size of SGBM is set to 5. The fourth is to derive the depth in the overlap region from

$$d = \frac{bfL}{pS} \tag{5}$$

where the parallax p is calculated by the matching algorithm, the projected sensor size S is 0.186 m, and the pixel length L is 1860. As a result, the calculated depth map alongside the ground truth is shown in Figure 13. The calculated depths of

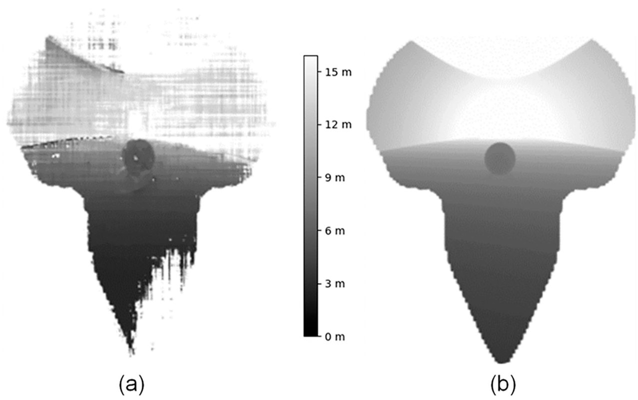


Figure 13. (a) Calculated depth map. (b) Ground truth. The calculated depths of most areas are close to the ground truth. However, in some areas, e.g., the lower right corner of the ground and parts of the wall and sky, the values of depth are missing or inaccurate.

most areas are close to the ground truth. However, in some areas, e.g., the lower right corner of the ground and parts of the wall and sky, the values of depth are missing or inaccurate.

5.4. Edge Map

Sharpening is an essential preprocessing step prior to detecting the edges, as it would have a dramatic impact on the results. To do so, a sharpening technique called unsharp masking is carried out. As for the Canny edge detection, its low and high thresholds are chosen as 60 and 120, respectively. The simulation results of edge detection are shown in Figure 14. Overall, the edge map with sharpening shows more details than the one without sharpening. For example, the unsharpened detection algorithm ignores most of the clouds in the sky and shadows on the ground, while a sharpened algorithm shows these edges.

https://doi.org/10.1021/acsaom.3c00431
ACS Appl. Opt. Mater. 2024, 2, 1269−1276

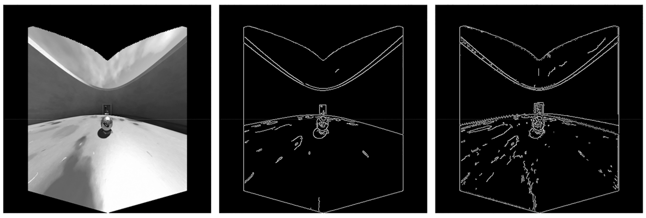


Figure 14. Edge map detected by the Canny algorithm. (a) Original bilateral neural image. (b) Edge map without sharpening. (c) Edge map with sharpening. Overall, the edge map with sharpening shows more details than the one without sharpening.

5.5. Saliency Map

Both visual saliency feature-based and spectral residual-based saliency detection algorithms are applied to draw the saliency map. For the visual saliency feature-based algorithm, as shown in Figure 15(a), the bright spot on the ball, shadow of the ball, clouds in the sky, and edges of the scene are marked as salient regions. For the spectral residual-based algorithm, as shown in Figure 15(b), the ball and edges of the scene are marked as salient regions. Given that the ball is arguably the only one meaningful object in the scene and it sits in the center, it shall be the most salient thing. For this reason, we tend to assume that the latter is more accurate, as the shadow of ball in the former does not make much sense. Incidentally, if the motion saliency detection is adopted and the input is a video, the static objects, e.g., the edges of scene, could be excluded.[39]

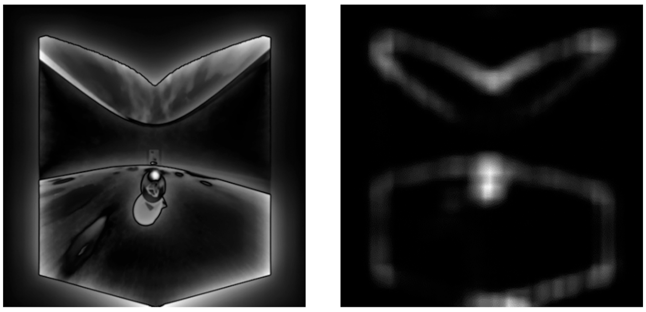


Figure 15. Saliency maps detected by the (a) visual saliency feature-based and (b) spectral residual-based algorithms, where the saliency is represented by the white color.

6. CONCLUSIONS

An epiretinal implant-based metaverse in conjunction with bionic vision processing has been conceptually probed. To sum up, we condense this work into the following contributions. Contribution 1: herald a next chapter of metaverse hardware, i.e., the retinal implants, for the next-gen humanity. Contribution 2: provide a methodology to simulate the stepwise images along the visual pathway, which are the monocular retinal image, unilateral neural image, bilateral neural image, binocular parallax-induced depth map, edge map, and saliency map. Contribution 3: refine the foveated blurring by linking its kernel size to the density of midget cells. Admittedly, to supersede the conventional devices, implantable devices still have a long road ahead. Issues like the postoperative reactions and biological incompatibility remain to be addressed.[40] Nevertheless, we believe that this technology will become popularized someday and give us a better metaverse.

■ ASSOCIATED CONTENT

Data Availability Statement

Data underlying the results presented in this paper are not publicly available but may be obtained from the authors upon reasonable request.

■ AUTHOR INFORMATION

Corresponding Author

Chao Ping Chen − *Smart Display Lab, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*; orcid.org/0000-0002-4366-4304; Email: ccp@sjtu.edu.cn

Authors

Haiyang Hu − *Smart Display Lab, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

Gang Li − *School of Psychology and Neuroscience, University of Glasgow, Glasgow G12 8QQ, United Kingdom*

Ziming Jin − *Smart Display Lab, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

Qiang Chu − *Smart Display Lab, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

Baoen Han − *Smart Display Lab, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

Seak Pang Zou − *Smart Display Lab, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsaom.3c00431

Author Contributions

Haiyang Hu: Writing - original draft, formal analysis, investigation, coding, software. Chao Ping Chen: Conceptualization, methodology, formal analysis, resources, writing - review and editing, supervision, project administration, funding acquisition. Gang Li: Formal analysis. Ziming Jin: Coding, data curation. Qiang Chu: Visualization. Baoen Han: Investigation. Seak Pang Zou: Visualization. All authors read and approved the final manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China (61831015), Science and Technology Commission of Shanghai Municipality (19ZR1427200), and Natural Science Foundation of Chongqing Municipality (cstb2023 nsq-msx0465, cstc2021jcyj-msxmX1136).

■ REFERENCES

(1) Harari, Y. N. *Homo Deus: A Brief History of Tomorrow*; Harvill Secker: London, UK, 2016.

(2) Bajaj, R.; Dagnelie, G.; Handa, J. T.; Scott, A. W. Retinal implants for RP: an update on Argus II and others. https://www.aao.

org/eyenet/article/retinal-implants-for-rp (accessed 29 January 2024).
(3) Hollister, S.; Peters, J. The FDA will apparently let Elon Musk put a computer in a human's brain. https://www.theverge.com/2023/5/25/23738123/neuralink-elon-musk-human-trial-fda-approval (accessed 29 January 2024).
(4) Chen, C. P.; Zhou, L.; Ge, J.; Wu, Y.; Mi, L.; Wu, Y.; Yu, B.; Li, Y. Design of retinal projection displays enabling vision correction. *Opt. Express* **2017**, *25* (23), 28223−28235.
(5) Chen, J.; Mi, L.; Chen, C. P.; Liu, H.; Jiang, J.; Zhang, W. Design of foveated contact lens display for augmented reality. *Opt. Express* **2019**, *27* (26), 38204−38219.
(6) Chen, C. P.; Mi, L.; Zhang, W.; Ye, J.; Li, G. Waveguide-based near-eye display with dual-channel exit pupil expander. *Displays* **2021**, *67*, 101998.
(7) Chen, C. P.; Cui, Y.; Chen, Y.; Meng, S.; Sun, Y.; Mao, C.; Chu, Q. Near-eye display with a triple-channel waveguide for metaverse. *Opt. Express* **2022**, *30* (17), 31256−31266.
(8) Chen, C. P.; Ma, X.; Zou, S. P.; Liu, T.; Chu, Q.; Hu, H.; Cui, Y. Quad-channel waveguide-based near-eye display for metaverse. *Displays* **2024**, *81*, 102582.
(9) Ahmad, K. M.; Klug, K.; Herr, S.; Sterling, P.; Schein, S. Cell density ratios in a foveal patch in macaque retina. *Vis. Neurosci.* **2003**, *20* (2), 189−209.
(10) Zeng, Q.; Zhao, S.; Yang, H.; Zhang, Y.; Wu, T. Micro/nano technologies for high-density retinal implant. *Micromachines* **2019**, *10* (6), 419.
(11) Jang, J.; Kim, H.; Song, Y. M.; Park, J.-U. Implantation of electronic visual prosthesis for blindness restoration. *Opt. Mater. Express* **2019**, *9* (10), 3878−3894.
(12) Hirsch, J.; Curcio, C. A. The spatial resolution capacity of human foveal retina. *Vision Res.* **1989**, *29* (9), 1095−1101.
(13) Kim, U. S.; Mahroo, O. A.; Mollon, J. D.; Yu-Wai-Man, P. Retinal ganglion cells—diversity of cell types and clinical relevance. *Front. Neurol.* **2021**, *12*, 661938.
(14) Wikipedia. Retinal ganglion cell. https://en.wikipedia.org/wiki/Retinal_ganglion_cell (accessed 20 January 2024).
(15) Wu, Y.; Chen, C. P.; Mi, L.; Zhang, W.; Zhao, J.; Lu, Y.; Guo, W.; Yu, B.; Li, Y.; Maitlo, N. Design of retinal-projection-based near-eye display with contact lens. *Opt. Express* **2018**, *26* (9), 11553−11567.
(16) Dragoi, V.; Tsuchitani, C. Chapter 14: visual processing: eye and retina. https://nba.uth.tmc.edu/neuroscience/s2/chapter14.html (accessed 29 January 2024).
(17) Dragoi, V.; Tsuchitani, C. Chapter 15: visual processing: cortical pathways. https://nba.uth.tmc.edu/neuroscience/s2/chapter15.html (accessed 29 January 2024).
(18) Schmittwilken, L.; Maertens, M. Fixational eye movements enable robust edge detection. *J. Vision* **2022**, *22* (8), 5.
(19) Qian, N. Binocular disparity and the perception of depth. *Neuron* **1997**, *18* (3), 359−368.
(20) Li, G.; McGill, M.; Brewster, S.; Chen, C. P.; Anguera, J. A.; Gazzaley, A.; Pollick, F. Multimodal biosensing for vestibular network-based cybersickness detection. *IEEE J. Biomed. Health Inform.* **2022**, *26* (6), 2469−2480.
(21) Bhartiya, S.; Ariga, M.; Puthuran, G. V.; George, R. *Practical Perimetry*; Jaypee Brothers Medical Publishers: New Delhi, India, 2016.
(22) Wikipedia. Field of view. https://en.wikipedia.org/wiki/Field_of_view (accessed 29 January 2024).
(23) Wikipedia. Visual acuity. https://en.wikipedia.org/wiki/Visual_acuity (accessed 29 January 2024).
(24) Watson, A. B. A formula for human retinal ganglion cell receptive field density as a function of visual field location. *J. Vision* **2014**, *14* (7), 15.
(25) Xi, N.; Ye, J.; Chen, C. P.; Chu, Q.; Hu, H.; Zou, S. P. Implantable metaverse with retinal prostheses and bionic vision processing. *Opt. Express* **2023**, *31* (2), 1079−1091.
(26) Wikipedia. Binocular vision. https://en.wikipedia.org/wiki/Binocular_vision (accessed 29 January 2024).
(27) Wikipedia. Ocular dominance. https://en.wikipedia.org/wiki/Ocular_dominance (accessed 29 January 2024).
(28) Goldstein, E. B.; Brockmole, J. *Sensation and Perception*, 10th ed.; Cengage Learning: Singapore, 2016.
(29) Burton, H. E. The optics of Euclid. *J. Opt. Soc. Am.* **1945**, *35* (5), 357−372.
(30) Wikipedia. Visual cortex. https://en.wikipedia.org/wiki/Visual_cortex (accessed 29 January 2024).
(31) Maini, R.; Aggarwal, H. Study and comparison of various image edge detection techniques. *Int. J. Image Process.* **2009**, *3* (1), 1−11.
(32) Sun, R.; Lei, T.; Chen, Q.; Wang, Z.; Du, X.; Zhao, W.; Nandi, A. K. Survey of image edge detection. *Front. Sig. Proc.* **2022**, *2*, 826967.
(33) Joshi, M.; Vyas, A. Comparison of Canny edge detector with Sobel and Prewitt edge detector using different image formats. *Int. J. Eng. Res. Technol.* **2020**, *1*, 133−137.
(34) Wikipedia, V1 saliency hypothesis. https://en.wikipedia.org/wiki/V1_Saliency_Hypothesis (accessed 29 January 2024).
(35) Montabone, S.; Soto, A. Human detection using a mobile platform and novel features derived from a visual saliency mechanism. *Image Vis. Comput.* **2010**, *28* (3), 391−402.
(36) Hou, X.; Zhang, L. Saliency detection: A spectral residual approach. In *Conference proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition*; June 17−22, **2007**, Minneapolis, MN; IEEE, 2007; pp 2438−2445.
(37) Bass, M.; DeCusatis, C.; Enoch, J. M.; Lakshminarayanan, V.; Li, G.; MacDonald, C.; Mahajan, V. N.; Van Stryland, E. *Handbook of Optics Vol. III: Vision and Vision Optics*, 3rd ed.; McGraw Hill: New York, NY, 2009.
(38) Hirschmuller, H. Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30* (2), 328−341.
(39) Cong, R.; Lei, J.; Fu, H.; Cheng, M. M.; Lin, W.; Huang, Q. Review of visual saliency detection with comprehensive information. *IEEE Trans. Circuits Syst. Video Technol.* **2019**, *29* (10), 2941−2959.
(40) Yue, L.; Weiland, J. D.; Roska, B.; Humayun, M. S. Retinal stimulation strategies to restore vision: fundamentals and systems. *Prog. Retin. Eye Res.* **2016**, *53*, 21−47.

https://doi.org/10.1021/acsaom.3c00431
ACS Appl. Opt. Mater. 2024, 2, 1269−1276