2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Electrode Dropout Compensation in Visual Prostheses: An Optimal Object Placement Approach

Reham H. Elnabawy, Slim Abdennadher, Olaf Hellwich, and Seif Eldawlatly

Abstract— **Visual prostheses provide promising solution to the blind through partial restoration of their vision via electrical stimulation of the visual system. However, there are some challenges that hinder the ability of subjects implanted with visual prostheses to correctly identify an object. One of these challenges is electrode dropout; the malfunction of some electrodes resulting in consistently dark phosphenes. In this paper, we propose a dropout handling algorithm for better and faster identification of objects. In this algorithm, phosphenes representing the object are translated to another location within the same image that has the minimum number of dropouts. Using simulated prosthetic vision, experiments were conducted to test the efficacy of our proposed algorithm. Electrode dropout rates of 10%, 20% and 30% were examined. Our results demonstrate significant increase in the object recognition accuracy, reduction in the recognition time and increase in the recognition confidence level using the proposed approach compared to presenting the images without dropout handling.**

***Clinical Relevance*— These results demonstrate the utility of dropout handling in enhancing the perception of images in prosthetic vision.**

I. Introduction

Among many causes of blindness, diseases such as retinitis pigmentosa and aged-macular degeneration represent one of the major causes that mainly affect the photoreceptors leading to loss of vision [1]. In recent years, different types of visual prostheses, aka bionic eyes, have been developed to artificially induce visual percepts through electrical stimulation of the visual pathway [2]. Retinal implants, such as the Argus II device, represent the most successful example of visual prostheses that target functional parts of the retina beyond the damaged site [3]. In such system, a tiny video camera on the bridge of eyeglasses captures images and sends them to a video processing unit. This unit transforms the images into electrical signals and sends them back to an antenna on the glasses. The signals are then transferred wirelessly to electrodes implanted either epi-retinal or sub-retinal. In addition to retinal implants, other types of visual prostheses have been proposed including cortical and thalamic types [4, 5].

Prosthetic vision induced via visual prostheses consists of spots of light called phosphenes that represent the pixels of the visual field [2]. Thus, the vision perceived via a visual prosthesis has been reported to be different from normal vision such that the patients must be ready to learn a new "language" of sight. Such vision is characterized by very low spatial and radiometric resolutions. The reduced quality of the perceived image results in a number of challenges such as psychological challenges due to over expectations from the patients, who think that they will restore their natural vision after implantation [6], and the difficulty that faces the patients in the beginning of using the device in recognizing the objects [7].

One other challenge that reduces the quality of the perceived image in prosthetic vision is stimulation electrode dropout. This occurs due to the malfunction of electrodes post-implantation or being implanted in dead tissue, which causes permanent black spots in the location of the corresponding phosphene(s) [8]. This negatively affects the recognition and identification of objects since some of the key-points in the object may be aligned with the receptive field of the neurons stimulated by the dropped-out electrodes. The most frequently observed percentages of phosphenes' dropouts have been shown to range from 10% to 30% [9, 10].

In this paper, we propose an image processing approach to compensate for the effects of electrode dropout on the perceived image. In this approach, convolution between a bounding box of the object of interest and the phosphene grid is used to identify and translate the object of interest to a location within the image that fits the object and has minimum dropouts. This is performed after showing the actual location of the object of interest to help in accurately identifying its location. We examined the performance of the proposed approach on different groups of normally sighted subjects using simulated prosthetic vision. Despite the differences between simulated and actual prosthetic vision, where actual prosthetic vision tends to be more complex, we adopted a phosphene simulation strategy that mimics perceived images reported by visual prostheses users [3, 11]. The results indicate a significant enhancement in recognizing the objects of interest using the proposed dropout handling approach. This could eliminate the need for performing additional interventions to replace the malfunctioning electrode.

II. Methods

A. Image Pre-processing

To allow better and faster identification of the object's identity in any prosthetic vision scene in the presence of dropouts, we propose an approach that optimally translate the

R. H. Elnabawy is with the Digital Media Engineering and Technology Department, Faculty of Media Engineering and Technology, German University in Cairo, Cairo, Egypt (e-mail: reham.elnabawy@guc.edu.eg).

S. Abdennadher is with the Computer Science and Engineering Department, Faculty of Media Engineering and Technology, German University in Cairo, Cairo, Egypt, e-mail: slim.abdennadher@guc.edu.eg)

O. Hellwich is with the Computer Vision and Remote Sensing Department, Faculty of Technical Information and Microelectronics, Technical University Berlin, Berlin, Germany (e-mail: olaf.hellwich@tu-berlin.de).

S. Eldawlatly is with the Computer and Systems Engineering Department, Faculty of Engineering, Ain Shams University, and the Faculty of Media Engineering and Technology, German University in Cairo, Cairo, Egypt (e-mail: seldawlatly@eng.asu.edu.eg).

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

object in the scene to a new location that has minimum number of dropouts. In this approach, the input image undergoes pre-processing prior to applying the dropout handling mechanism. In the pre-processing, the input image is first resized to 32 × 32 pixels that represent the number of hypothesized stimulation electrodes. The image is then converted to grayscale. Next, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to the image to enhance the image contrast [12]. This is necessary to enhance the contrast of the test images for better presentation. A Wiener filter of size 3 × 3 is then applied to remove any noise in the image to allow better identification of the object [13]. Next, Otsu thresholding is applied to perform an automatic image thresholding of the image [14]. This algorithm aims to find the threshold value where the sum of foreground and background spreads is at its minimum. This is followed by applying median filter of size 3 × 3 to the image to remove any remaining noise while preserving the image's details. Connected components labeling using 8-connectivity is then used to help in better object identification, where any region that is of an area with a negligible size is then removed [15].

B. Phosphene Simulation

The output of the pre-processing stages is used to generate a simulated prosthetic vision image. We used a square grid for phosphene representation [3]. A circular phosphene shape is also used in this study; consistent with multiple simulated prosthetic vision studies [11]. The distance between each two consecutive phosphenes was set to zero. Dropouts were then simulated with rates of 10%, 20% or 30% of the total number of pixels [10]. The location of the dropped out phosphenes was determined randomly following a uniform distribution. In all simulations, a dropped out phosphene was set to a black color.

C. Dropout Handling

The main aim of this paper is to minimize the impact of dropouts on the perception. This is done by translating the object in a certain image to a certain place in the same image where the number of dropped out phosphenes is minimum. In this approach, we construct a matrix D of size 32 × 32 (i.e., the same enhanced input image size) in which dropouts are represented by 0, while other pixels are represented by 1. A bounding box B is then identified around the object of interest by computing the connected-components labeling using 8-connectiviy to get the connected pixels that refer to the same object and retrieve the minimum row, the minimum column, the maximum row and the maximum column. All pixels within B are set to 1. A convolution process is then performed between B and D, and the optimal location for the center of the object of interest C_{New} is identified by the position that has the highest value in the convolution output defined as

$$C_{New} = max_{(x,y)}\{(B * D)[x, y]\} \tag{1}$$

Finally, we translate the object within B in the presented image to C_{New}. If multiple locations have the same maximum value in the convolution output, C_{New} is set as the location that is closest to the center of B based on the Euclidean distance.

D. Experimental Design and Procedure

To examine the efficacy of the proposed approach, 12 subjects participated in the experiments (5 males and 7 females) of age 22 to 60 years. All the participants in the experiments had normal/corrected vision. Simulated test images were presented to the subjects using a 15-inch computer screen. Subjects were seated at a distance of 1 meter from the computer screen. This resulted in a visual field angle of 20° which mimics the visual filed of visual prostheses [16]. Prior to the presentation of the images, each subject was given a demonstration that included 3 different test images that are different from the images used in the following experiments to avoid any learning effects. Two versions of the same image were used in the demonstration: The first version is displayed in terms of phosphenes to introduce the subjects to prosthetic vision, and the second version demonstrated the effect of phosphene dropout (i.e., black spots).

The test subjects were divided into 4 groups with 3 subjects each. Each group participated in a different experiment in which each subject was presented with 24 different test images: 8 images with 10% dropout, 8 images with 20% dropout and 8 images with 30% dropout. The 24 test images set was fixed across all subjects. Each test image was displayed in a trial of duration 10 sec and then the subjects were given the chance to tell the identity of the object displayed in the image. The images represented objects from different categories including car, utensils, flower, window, bed, chair, numbers, bird, teapot, stairs, shelf, and truck.

The first group of subjects was presented with the test images after performing the pre-processing and prosthetic vision simulation for the whole 10 sec. No dropout handling was performed. This group represents what current visual prostheses users would perceive. The second group was presented with the same image presented to the first group for 5 sec only. The object is then translated to a random position within the image and displayed for another 5 sec. This test was performed to determine whether any enhancement in performance was in fact due to the proposed dropout handling approach or a result of displaying two different versions of the object to the subject. Dropout handling was examined in the two other groups. The third group was presented with the output of the proposed dropout handling approach for the entire 10 sec. The last group was presented with the test images after pre-processing and prosthetic vision simulation without dropout handling for 5 sec, and then with the output of the dropout handling approach for another 5 sec. This last group represents, contrary to the third group, a practical implementation of the proposed approach as it still provides the subject with the actual location of the object of interest during the first 5 sec, and then a better presentation of the object for another 5 sec after dropout handling. This could help in providing the subject with an accurate localization of the object while minimizing the effects of dropout.

E. Evaluation Metrics

Three metrics were used to assess the performance of each group. First, we measured the object recognition accuracy as

$$Accuracy\,\% = \frac{N}{T} \times 100 \tag{2}$$

where N is the number of correctly identified objects by the subject and T is the total number of presented test images. Second, the time taken by the subject to correctly recognize the presented object was recorded, with a maximum of 10 sec. Finally, the subjects reported their confidence level in the recognition on a scale of 1 to 5.

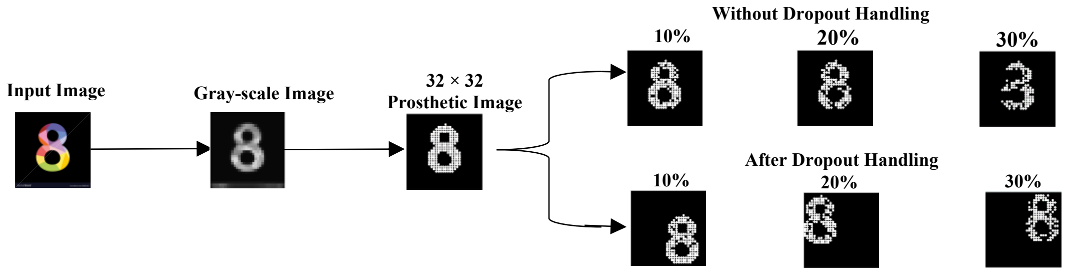

Figure 1. Dropout simulation and handling for different dropout rates of 10%, 20% and 30%.

III. Results

A. Dropout Simulation and Handling

We first demonstrate the output of each stage of the proposed approach. Fig. 1 illustrates an overview of the process implemented to show the effect of the dropout handling approach on the object clarification and, thus, object recognition. It shows a sample test image representing the number "eight" for different dropout rates. The figure demonstrates, especially in the case of 30% dropout rate, how translating the object of interest to a location with minimum number of dropouts enhances the presentation.

B. Performance Evaluation

We examined the utility of the proposed approach by presenting different sets of images to four groups of subjects. For each group, images with 10%, 20% and 30% dropout rates were presented. Fig. 2 illustrates the performance of the test subjects for a dropout rate of 10%. Fig. 2a demonstrates significantly higher recognition accuracy when the proposed dropout handling approach was used compared to not applying dropout handling and compared to presenting the subjects with the images without dropout handling for 5 sec in addition to a randomly translated version of the image for another 5 sec (No dropout handling: 66.67 ± 29.56%, Random placement: 77.1 ± 19.79%, Dropout handling: 97.91 ± 5.9%, $P < 0.05$, $n = 24$, two-sample *t*-test). Moreover, the figures demonstrate that dropout handling results in significantly shorter recognition time (Fig. 2b; No dropout handling: 8.68 ± 1.55 sec, Random placement: 7.88 ± 2.38 sec, Dropout handling: 3.33 ± 1.28 sec, $P <$ 1e-07, $n = 24$, two-sample *t*-test) and higher decision confidence (Fig. 2c; No dropout handling: 3.48 ± 0.79, Random placement: 3.75 ± 0.74, Dropout handling: 4.91 ± 0.24, $P <$ 1e-04, $n = 24$, two-sample *t*-test).

One disadvantage of applying the proposed approach is that it modifies the actual location of the displayed object. This might be confusing to a visual prosthesis user as it does not accurately represent the visual field of the subject. Therefore, to provide the subjects with both the actual location of the object in addition to a better representation with least dropouts, we modified the image presentation for the last group of subjects. In this presentation, the image is presented without dropout handling for 5 sec followed by presenting the image after dropout handling for another 5 sec. A significantly better performance can still be observed using this approach compared to no dropout handling and the random placement approaches (Recognition accuracy: 97.9 ± 5.9%, Time: 6.24 ± 0.94 sec, Confidence level: 4.84 ± 0.35, $P < 0.05$, $n = 24$, two-sample *t*-test). In addition, no significant difference can be observed between this modified presentation and applying the dropout handling for the entire 10 sec in terms of both the

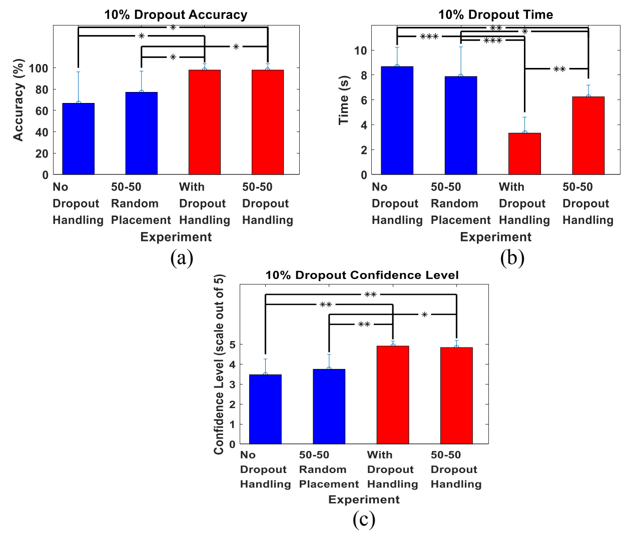

Figure 2. Experiments statistics analysis for 10% dropout rate for three metrics (a) recognition accuracy, (b) time to decision, and (c) confidence level (mean ± std). Blue bars represent not using dropout handling, while red bars represent two different versions of dropout handling. $*P < 0.05$, $**P <$ 1e-04, $***P <$ 1e-07, two-sample *t*-test.

recognition accuracy and the confidence level. However, in terms of the time taken to decide the identity of the object, higher time was needed by the subjects. This is expected given that for half of the time (i.e., 5 sec), the dropout handling approach was not applied, which contributes to the time needed to recognize the object.

Examining the performance of the four groups of subjects with dropout rates of 20% and 30% as demonstrated in Fig. 3 and Fig. 4, respectively, revealed a consistent enhancement when using the proposed dropout handling approach. First, significantly higher recognition accuracy can be observed using the dropout handling approach applied for the entire 10 sec compared to not applying dropout handling and the random placement approaches in the case of 20% dropout rate (No dropout handling: 62.5 ± 44.32%, Random placement: 58.34 ± 41.79%, Dropout handling: 95.83 ± 7.73%, $P < 0.05$, $n = 24$, two-sample *t*-test). For the 30% dropout rate, higher, yet not statistically significant, recognition accuracy can be observed using dropout handling (No dropout handling: 87.5 ± 35.35%, Random placement: 83.34 ± 35.63%, Dropout handling: 100%). The lack of significance in this case can be explained given that for 30% dropout rate, relatively simpler test images were used to compensate for the high dropout rate used. Second, enhancement in the performance can be observed in both the time taken to decide the identity of the presented object and the decision confidence for a dropout rate of 20% (Time: No dropout handling: 8.93 ± 1.51 sec, Random placement: 8.41 ± 2.22 sec, Dropout handling: 5.08 ± 1.69 sec,

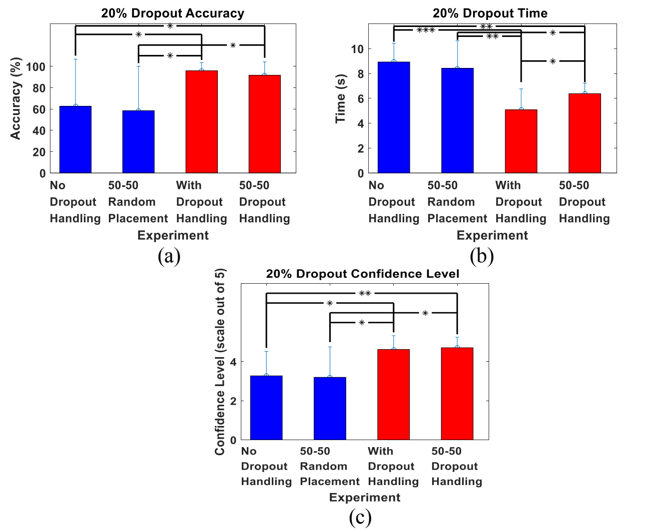

Figure 3. Experiments statistics analysis for 20% dropout rate for three metrics (a) recognition accuracy, (b) time to decision, and (c) confidence level (mean ± std). Blue bars represent not using dropout handling, while red bars represent two different versions of dropout handling. **P* < 0.05, ***P* < 1e-04, ****P* < 1e-07, two-sample *t*-test.

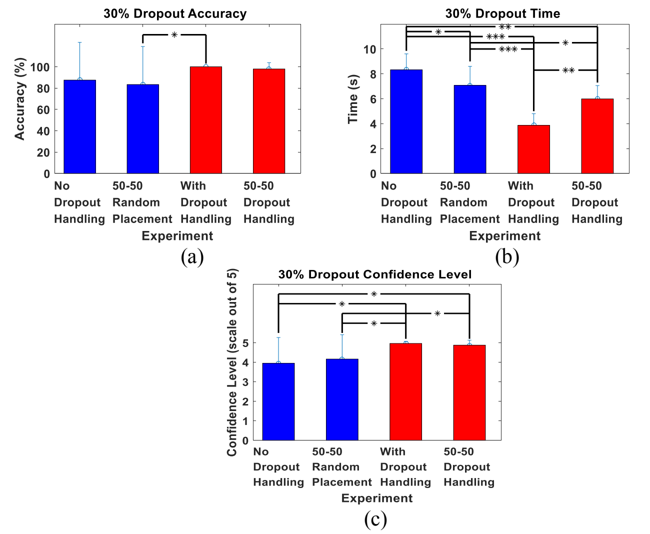

Figure 4. Experiments statistics analysis for 30% dropout rate for three metrics (a) recognition accuracy, (b) time to decision, and (c) confidence level (mean ± std). Blue bars represent not using dropout handling, while red bars represent two different versions of dropout handling. **P* < 0.05, ***P* < 1e-04, ****P* < 1e-07, two-sample *t*-test.

P < 1e-07, n = 24, two-sample t-test; Confidence level: No dropout handling: 3.28 ± 1.25, Random placement: 3.2 ± 1.6, Dropout handling: 4.63 ± 0.7, P < 1e-04, n = 24, two-sample t-test) and a dropout rate of 30% (Time: No dropout handling: 8.33 ± 1.28 sec, Random placement: 7.08 ± 1.53 sec, Dropout handling: 3.86 ± 0.93 sec, P < 1e-07, n = 24, two-sample t-test; Confidence level: No dropout handling: 3.95 ± 1.32, Random placement: 4.16 ± 1.24, Dropout handling: 4.96 ± 0.11, P < 1e-04, n = 24, two-sample t-test). Finally, applying the dropout handling approach for 5 sec only out of the entire 10 sec showed no significant difference in both the recognition accuracy and the confidence level compared to using the dropout handling approach for the entire 10 sec for both 20% and 30% dropout rates. However, similar to the 10% dropout rate, significant increase in the time taken to decide the identity of the object can be observed.

IV. Conclusion

Visual prostheses have recently demonstrated success in restoring vision to the blind. One of the challenges that affect the quality of the perceived prosthetic image is electrode dropout. We proposed an approach that could help in better recognition of the identity of an object by means of convolution and translation. The object's phosphenes are optimally translated to a location with minimum number of electrodes dropouts. Experiments using simulated prosthetic vision revealed a significant enhancement in the ability of the test subjects to recognize the presented objects using the proposed approach compared to presenting the test image without dropout handling to the subjects as well as randomly translating the object within the image. Additionally, for the practical utilization of the approach, presenting the output of the proposed approach to the test subjects subsequent to presenting the image without dropout handling showed similar enhancement. These results indicate the efficacy of the proposed approach in compensating for the effects of electrode dropout in visual prostheses. Further studies that involve visual prostheses users should be pursued to better assess the utility of the proposed approach by measuring the accuracy and time taken for object identification and localization.

References

[1] F. Newton and R. Megaw, "Mechanisms of Photoreceptor Death in Retinitis Pigmentosa," *Genes,* vol. 11, p. 1120, 2020.
[2] G. Dagnelie, *Visual prosthetics: physiology, bioengineering, rehabilitation*: Springer Science & Business Media, 2011.
[3] Y. H.-L. Luo and L. Da Cruz, "The Argus® II retinal prosthesis system," *Progress in retinal and eye research,* vol. 50, pp. 89-107, 2016.
[4] P. M. Lewis*, et al.*, "Restoration of vision in blind individuals using bionic devices: a review with a focus on cortical visual prostheses," *Brain research,* vol. 1595, pp. 51-73, 2015.
[5] J. S. Pezaris and R. C. Reid, "Demonstration of artificial visual percepts generated through thalamic microstimulation," *Proceedings of the National Academy of Sciences,* vol. 104, pp. 7670-7675, 2007.
[6] H. C. Stronks and G. Dagnelie, "The functional performance of the Argus II retinal prosthesis," *Expert review of medical devices,* vol. 11, pp. 23-30, 2014.
[7] A. P. Finn*, et al.*, "Argus II retinal prosthesis system: a review of patient selection criteria, surgical considerations, and post-operative outcomes," *Clinical Ophthalmology (Auckland, NZ),* vol. 12, p. 1089, 2018.
[8] Y. Luo, "Argus® II Retinal Prosthesis System: Clinical & Functional Outcomes," UCL (University College London), 2017.
[9] G. Dagnelie*, et al.*, "Real and virtual mobility performance in simulated prosthetic vision," *Journal of neural engineering,* vol. 4, p. S92, 2007.
[10] Y. Zhao*, et al.*, "Chinese character recognition using simulated phosphene maps," *Investigative ophthalmology & visual science,* vol. 52, pp. 3404-3412, 2011.
[11] S. C. Chen*, et al.*, "Simulating prosthetic vision: I. Visual models of phosphenes," *Vision research,* vol. 49, pp. 1493-1506, 2009.
[12] R. Manju*, et al.*, "Improved Method for Enhancing Dark Images based on CLAHE and Morphological Reconstruction," *Procedia Computer Science,* vol. 165, pp. 391-398, 2019.
[13] S. Kumar*, et al.*, "Performance comparison of median and wiener filter in image de-noising," *International Journal of Computer Applications,* vol. 12, pp. 27-31, 2010.
[14] J. Zhang and J. Hu, "Image segmentation based on 2D Otsu method with histogram analysis," in *2008 international conference on computer science and software engineering*, 2008, pp. 105-108.
[15] L. He*, et al.*, "The connected-component labeling problem: A review of state-of-the-art algorithms," *Pattern Recognition,* vol. 70, pp. 25-43, 2017.
[16] M. Sanchez-Garcia*, et al.*, "Semantic and structural image segmentation for prosthetic vision," *PloS one,* vol. 15, p. e0227677, 2020.