35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Embracing the Irregular: A Patient-Specific Image Processing Strategy for Visual Prostheses

Filiz Isabell Kiral-Kornek*[,1,2], Craig O. Savage[1,2], Elma O'Sullivan-Greene[1,2], Anthony N. Burkitt[1,2,3,4], and David B. Grayden[1,2,3,4]

Abstract— **We propose a stimulation strategy for retinal prostheses that makes use of irregular shapes of elicited phosphenes. It is patient specific and thus relies on prior psychophysical measurements. Visual perceptions are stored in a phosphene map that relates stimulation parameters to the visual stimulus elicited. Based on this map, stimulation parameters are chosen in such a way that the edges of the target image are optimally represented through the shape of the phosphene. In a psychophysical pilot study, we compare this approach to one in which we choose phosphenes to match the brightness of the target image. We find that participants perform similarly well with both strategies overall. However, the results indicate that each strategy may have advantages for different stimulus sizes. Both of the proposed strategies are novel in using only previously recorded phosphenes rather than a model based on idealized assumptions about the relationship between stimulation parameters and phosphene properties.**

I. INTRODUCTION

Research groups around the world are working on building visual prostheses to restore sight for people suffering from degenerative retinal diseases such as retinitis pigmentosa and age-related macular degeneration [6, chapter 1]. A characteristic of both diseases is the degeneration of photoreceptors in the retina. Even after these light sensitive cells have died, the remaining cellular network can still be intact. To make use of the remaining retina, a common approach is to electrically stimulate retinal ganglion cells [6, chapter 6]. In recent projects, a camera mounted on glasses records the scene in front of the patient and sends the information to an encoding unit that transforms the image signal into electrical pulses. These are sent to an electrode array inside the eye that can sit on or under the retina [5], [13] or behind the choroid [12]. An electric current is then supplied to each individual electrode with the intention to evoke a visual perception for each stimulus. The elicited perceptions, called phosphenes, are commonly modeled as bright spots with the brightness decay following a two-dimensional Gaussian distribution. Brightness is assumed to be proportional to the current density delivered to the tissue under the corresponding electrode.

Using this model of visual perception, stimulation strategies are being developed, most of them proposing to use the brightness of phosphenes to represent the brightness of the corresponding part of the scene in front of the patient. Alternatively, a depth mapping has been proposed, presenting objects that are closer to the patient with brighter phosphenes than objects that are farther away [1]. A third approach is to accentuate objects that are considered to be more important, such as faces or tripping hazards, with brighter phosphenes [10].

However, psychophysical measurements conducted with implanted patients have revealed that this phosphene model is not particularly realistic. Phosphenes are rarely round, but can instead have a wide range of irregular shapes and sometimes even different colors [11]. Some patients have reported different brightness levels within one elicited phosphene. This can make it difficult to assign a specific brightness level to a phosphene, which in turn complicates the implementation of the stimulation strategies mentioned above.

There may, however, be a way to use the shape of phosphenes rather than their brightness. By composing an image out of the differently shaped phosphenes as they are perceived by patients, we present an alternative strategy to communicate visual information.

In the case that we are studying, we depart from the assumption that we can influence brightness directly with current and only use phosphenes that have previously been reported by the implantee. Because these phosphenes were shown to be irregular at times and even have different brightnesses within one phosphene, choosing a phosphene of a specific brightness level can be difficult.

The remainder of this paper is structured as follows: In Section II, we present a phosphene map which describes the relationship between physical stimulation parameters and the elicited visual perception in Section II-A. Section II-B explains an approach to match the brightness of irregular phosphenes to the brightness of an input image. Section II-C demonstrates a correlation-based approach to pick the phosphene out of the phosphene map that is closest in shape to the image part that we want to represent. Both stimulation strategies are tested on normally sighted observers in psychophysical simulations, which are described in Section II-D. Results of these tests are given in III. The two approaches are then compared and discussed in Section IV. Section V concludes this document.

[1]NeuroEngineering Laboratory, Department of Electrical & Electronic Engineering, The University of Melbourne VIC, 3010, Australia
[2]Centre for Neural Engineering, The University of Melbourne
[3]NICTA, c/- Department of Electrical & Electronic Engineering, The University of Melbourne, Victoria, 3010, Australia
[4]Bionics Institute, 384-388 Albert St, East Melbourne VIC 3002, Australia
*Correspondence to be sent to:
i.kiral-kornek@student.unimelb.edu.au

II. METHODS

A. Phosphene Map

For our application, a phosphene map describes the relationship between the perception of the patient and stimulus parameters such as the frequency and amplitude of the applied stimulation current and the position of the electrode. Phosphenes are taken from various sources in the literature of reported psychophysics tests [11], [7], [8], [4]. Figure 1 displays available phosphene shapes and brightnesses used. Additionally, the off-state of an electrode is assumed to result in no perception or a black phosphene. To create a mapping between stimulation parameters and the perceived phosphene, we make the assumption that every electrode can elicit the same range of phosphenes. This simplification has to be made in lieu of complete patient data. We also assume that each phosphene is evoked by stimulating one single electrode. At this stage, interactions between electrodes are not considered.

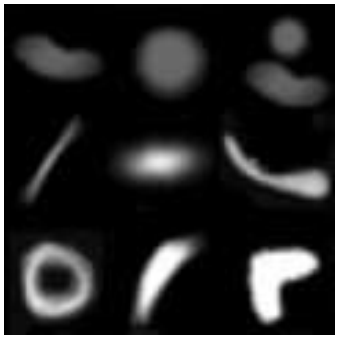

Fig. 1. All phosphenes represented in the phosphene map.

Phosphenes are stored as 20x20 pixel monochromatic images. Each phosphene has coordinates that are defined in relation to the electrode position. To get a more realistic representation, we apply some random local jitter to each phosphene, but limited to an extent to which phosphenes associated with different electrodes don't overlap. Brightness and shape of phosphenes are described by phosphene images that are stored in the JPG format.

If this method were to be implemented in a real device, the perception would need to be measured separately for each set of stimulation parameters, including the simultaneous stimulation of different electrodes.

In the following paragraphs, two methods of composing an image will be described. Rather than assuming a perfect grid and ideal phosphenes, both methods rely on the perceptions stored in the phosphene map.

B. Using Phosphene Brightness

In the phosphene brightness stimulation strategy, we calculate the average brightness $\overline{P_{e,p}}$ of each phosphene p for each electrode e and normalize this value over the entire range of available phosphenes. From all phosphenes that can be evoked by one electrode, we then compare the brightness at its specific location to the brightness of the image segment that we want to display and choose the one with the smallest difference.

In a first step, the input image that we want to display is low-pass filtered with a two-dimensional Gaussian. This input image can either be a gray scale image, a depth map, or even an edge representation of the scene. For each phosphene of each associated electrode, the normalized brightness of the image segment in the area of the field of regard where a phosphene would be perceived is calculated as:

$$\overline{I_{e,p,norm}} = \text{mean}\left(\sum_{x=c_x-x_r/2}^{c_x+x_r/2}\sum_{y=c_y-y_r/2}^{c_y+y_r/2}((I(x,y)-\overline{I}_{min})/\overline{I}_{max})\right), \quad (1)$$

where c_x and c_y as an electrode position and x_r and y_r define the range of the image segment in pixels. The minimum and maximum brightness values of all image segments are subsequently denoted $\overline{I}_{min}$ and $\overline{I}_{max}$ subsequently.

The brightness of each phosphene is calculated similarly by:

$$\overline{P_{e,p,norm}} = \text{mean}\left(\sum_{x=c_x-x_r/2}^{c_x+x_r/2}\sum_{y=c_y-y_r/2}^{c_y+y_r/2}((P(x,y)-\overline{P}_{min})/\overline{P}_{max})\right). \quad (2)$$

For each electrode, the phosphene that results in the minimum absolute value of $\overline{I_{e,p,norm}} - \overline{P_{e,p,norm}}$ is presented to the implantee. Figures 2(b) and 2(f) show examples of the output for the letters "Z" and "S", respectively.

C. Using Phosphene Shape

In the case of more complex phosphenes that can be composed of multiple percepts with different brightnesses, it is difficult to determine one exact brightness level for each phosphene. As an alternative to the strategy presented in II-B, we choose phosphenes based solely on the shape as stored in the phosphene map. To obtain the best representation of an input image, we first detect the edges in the image using a canny edge detector [3]. We then compute the two-dimensional cross-correlation of each phosphene with the image segment at the position of that phosphene. For each electrode, the phosphene with the highest cross-correlation is chosen for presentation to the implantee.

For a phosphene with the brightness $P(x,y)$ at the positsion (x,y) representing an image segment with brightness $I(x,y)$, the cross-correlation is given by:

$$C(i,j) = \sum_{x=0}^{c_x+x_r}\sum_{y=0}^{c_y+y_r} P(x,y)\cdot \text{conj}(I(x+i,y+j)), \quad (3)$$

where conj denotes the conjugate. Because the cross-correlation is maximal when two shapes align perfectly, we find the maximum value for each phosphene electrode pair and choose the phosphene with the maximal value.

Figures 2(c) and 2(g) show examples of the output for the letters "Z" and "S" using this cross-correlation based method, respectively.

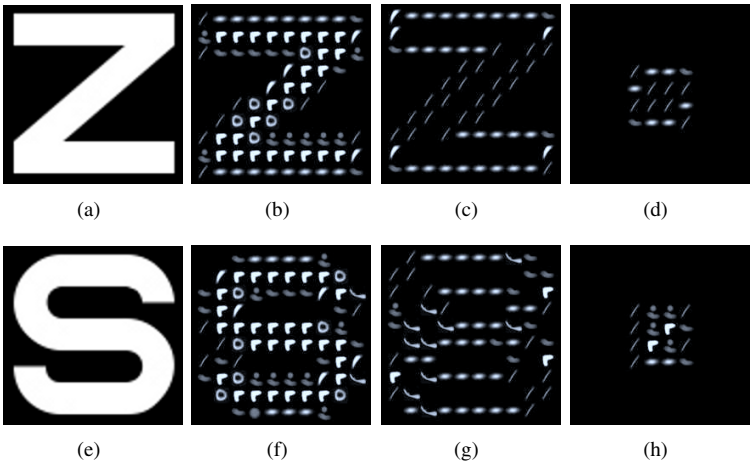

Fig. 2. Input letter "Z" (a), using brightness matching to represent letter "Z" (b), using shape matching to represent letter "Z" (c), using shape matching to represent letter "Z" in reduced size (d), input letter "S" (e), using brightness matching to represent letter "S" (f), using shape matching to represent letter "S" (g), using brightness matching to represent letter "Z" in reduced size (h).

D. Psychophysical Measurement

To compare both strategies, we measured the percentage of correctly recognized stimuli in a psychophysical pilot study. The study was conducted in agreement with the Australian National Statement Ethical Conduct and was approved by the Human Research Ethics Committee of the University of Melbourne.

Six participants with normal or lens corrected vision were asked to distinguish between ten different letters (C, D, H, K, N, O, R, S, V, Z) as defined in the British Standard 4274-1:2003 [2] that were presented using either of the two proposed strategies. The experiment was conducted as a 10 alternative forced choice test. One half of the participants started with the brightness-based, the other half with the shape-based stimulation strategy.

The independent variable was the stimulus size. Examples of letters of different sizes can be found in Figure 2. All other properties were left unchanged. As dependent variable, we measured the percentage of correctly recognized letters.

Stimuli were created with Matlab and presented on a Dell UltraSharp U2312HM 23inch LED monitor. A chin rest was used to fix the participant's head position relative to the screen at a distance of 57 cm. Participants initially underwent a short training session to get used to the task and stimuli. Data were collected at stimulus sizes $0.15, 0.3, 0.4, 0.5$, and $0.65S_I$, with S_I taken as the size of the input image which corresponds with 5 cm on the screen or 5 degrees of the visual angle of the participant. First, a white noise mask image was displayed for 500 ms to prepare the participant for the stimulus, followed by the stimulus for 160 ms, and another mask image for 500 ms to prevent after-images. The stimulus time was chosen in a way that made scanning the image impossible for the participant. The participant then had to name the letter they perceived. During this time, the screen was blank. The experimenter typed in the response and the next stimulus was presented. Each participant was presented with 200 stimulus images per strategy with an equally distributed letters.

III. RESULTS

Figure 3 displays the results of the psychophysical experiment. The graph shows the fraction of correctly identified letters against the size of the stimulus for both stimulation strategies described above. Blue stars represent the mean over all experiments using the brightness matching strategy, red diamonds represent the mean fraction of correctly identifies letters using the shape matching algorithm. The standard deviation across all experiments is shown by error bars around the mean values. A sigmoidal distribution (logistic function) is fitted to the mean fraction of correctly identified letters for both stimulation strategies.

IV. DISCUSSION

While most stimulation strategies rely on idealized inferred relationships between stimulation parameters and the visual appearance of a phosphene, our approaches rely on a previously measured phosphene map that only includes the actual percepts that patients describe. There are of course difficulties in acquiring such a phosphene map, especially for prostheses with large numbers of electrodes. Each set of stimulation parameters has to be measured separately, a rather time-consuming undertaking, which gets more extensive once simultaneous stimulation of electrodes is included.

However, if the deliberate manipulation of phosphene shape becomes possible as for example suggested in [9], for larger arrays it may be an option to stimulate electrodes in smaller groups to evoke different shapes.

Preliminary data presented in Section III appears to indicate for letter representation that it can be beneficial to use a shape matching strategy while for larger stimulation sizes a brightness matching strategy may be favorable.

A possible extension to this work is expand the available phosphene database by implementing electrode specific phosphenes with random shapes assigned to each electrode to ensure a more realistic variety of representation. In a last

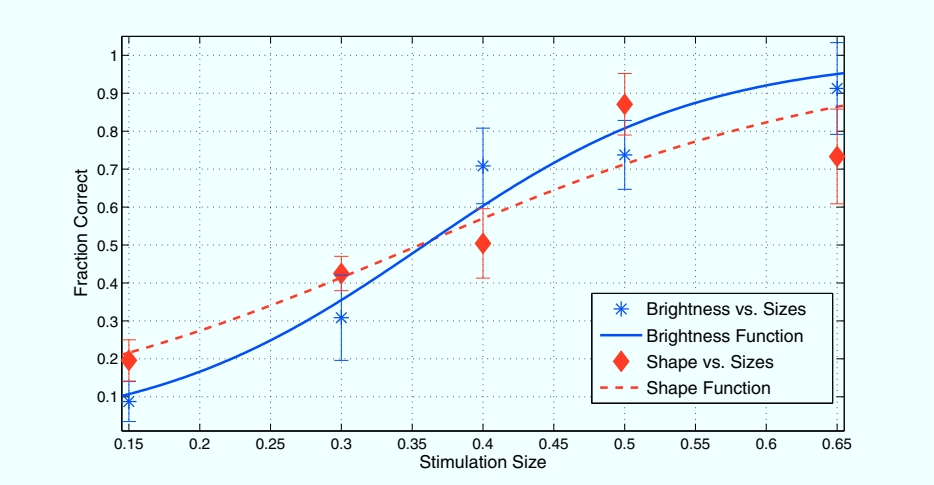

Fig. 3. Fraction of correctly identified letters for all experiments over stimulation size, using brightness (blue) and shape (red) matching.

extension an actual patient-specific phosphene map can be used. This would enable us to compare our strategy to the output of a strategy that relies on idealized relationships between stimulation parameters and visual percepts.

V. CONCLUSION

In this paper, two novel stimulation strategies are proposed for a prosthesis visual device. These strategies are solely based on a database of previously measured visual percepts that are stored in a phosphene map with their corresponding stimulation parameters. While existing strategies are based on highly idealized models relating stimulation parameters to symmetric phosphenes of uniform brightness, instead, we leverage upon the highly irregular and patient specific nature of true phosphene elicitation to provide more optimal and targeted information transfer.

By conducting a letter recognition experiment, we showed that using true phosphene shapes to represent letters is feasible in practice with reasonable letter identification achieved. Furthermore, rather than being problematic, the non-idealities in real phosphenes can be used beneficially by capitalizing on directional information to convey fine detail in text to a visual prosthesis user. Larger studies are required to determine if this holds for more complex stimuli. However, it could be envisioned for future devices that image processing could highlight other directional information, for example, to emphasize edges in an image to help with navigation.

ACKNOWLEDGMENT

This research was supported by the Australian Research Council (ARC) through its Special Research Initiative (SRI) in Bionic Vision Science and Technology grant to Bionic Vision Australia (BVA). The Bionics Institute acknowledges the support it receives from the Victorian Government through its Operational Infrastructure Support Program.

REFERENCES

[1] N. Barnes, P. Lieby, H. Dennet, J. Walker, C. McCarthy, N. Liu, and Y. Li, "Investigating the role of single-viewpoint depth data in visually-guided mobility," *Journal of Vision*, vol. 11, no. 11, pp. 926–926, 2011.

[2] British Standards Institute, "BS 4274-1: Visual acuity test types. test charts for clinical determination of distance visual acuity. specification." British Standards, 2003.

[3] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6, pp. 679–698, 1986.

[4] A. Horsager and I. Fine, "The perceptual effects of chronic retinal stimulation," *Visual prosthetics: physiology, bioengineering, rehabilitation. New York, NY: Springer*, pp. 271–300, 2011.

[5] M. Humayun, J. Weiland, G. Fujii, R. Greenberg, R. Williamson, J. Little, B. Mech, V. Cimmarusti, G. Van Boemel, G. Dagnelie *et al.*, "Visual perception in a blind subject with a chronic microelectronic retinal prosthesis," *Vision research*, vol. 43, no. 24, pp. 2573–2581, 2003.

[6] J. Martins and L. Sousa, *Bioelectronic vision*. World Scientific, 2009.

[7] D. Nanduri, I. Fine, A. Horsager, G. Boynton, M. Humayun, R. Greenberg, and J. Weiland, "Frequency and amplitude modulation have different effects on the percepts elicited by retinal stimulation," *Investigative Ophthalmology & Visual Science*, vol. 53, no. 1, pp. 205–214, 2012.

[8] D. Nanduri, M. Humayun, R. Greenberg, M. McMahon, and J. Weiland, "Retinal prosthesis phosphene shape analysis," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. IEEE, 2008, pp. 1785–1788.

[9] C. Savage, F. Kiral-Kornek, B. Tahayori, and D. Grayden, "Can electric current steering be used to control perception of a retinal prosthesis patient," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 3013–3016.

[10] A. Stacey, Y. Li, and N. Barnes, "A salient information processing system for bionic eye with application to obstacle avoidance," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011, pp. 5116–5119.

[11] R. Wilke, V. Gabel, H. Sachs, K. Schmidt, F. Gekeler, D. Besch, P. Szurman, A. Stett, B. Wilhelm, T. Peters *et al.*, "Spatial resolution and perception of patterns mediated by a subretinal 16-electrode array in patients blinded by hereditary retinal dystrophies," *Investigative ophthalmology & visual science*, vol. 52, no. 8, pp. 5995–6003, 2011.

[12] J. Zhou, S. Woo, S. Park, E. Kim, J. Seo, H. Chung, and S. Kim, "A suprachoroidal electrical retinal stimulator design for long-term animal experiments and in vivo assessment of its feasibility and biocompatibility in rabbits," *Journal of Biomedicine and Biotechnology*, vol. 2008, 2008.

[13] E. Zrenner, K. Bartz-Schmidt, H. Benav, D. Besch, A. Bruckmann, V. Gabel, F. Gekeler, U. Greppmaier, A. Harscher, S. Kibbel *et al.*, "Subretinal electronic chips allow blind patients to read letters and combine them to words," *Proceedings of the Royal Society B: Biological Sciences*, vol. 278, no. 1711, pp. 1489–1497, 2011.