AN OVERVIEW OF VISION PROCESSING IN IMPLANTABLE PROSTHETIC VISION

Nick Barnes

Computer Vision Research Group, NICTA, The Australian National University, Bionic Vision Australia

ABSTRACT

Electrically stimulating prosthetic vision devices offer a potential therapy to blind individuals. There are currently two multi-centre trials of devices by Second Sight Medical Products, and by Zrenner's group at University of Tuebingen. In Australia, Bionic Vision Australia has a retinal implant trial with three patients. Current implants provide restricted information for implantees, and some limitations are likely to remain in the future. To provide a substantial benefit to individual's abilities to perform key tasks such as orientation and mobility, activities of daily living, reading and face recognition there is much work to be done. Vision processing's role is to ensure the key visual information is available to undertake tasks given these limitations. This paper frames the background and challenges in vision processing for implantable prosthetic vision, and gives an overview of recent work.

Index Terms— Visual prosthesis, Vision processing for implantable prosthetic vision, blindness, retinal implants.

1. INTRODUCTION

In this paper, we introduce vision processing for implantable prosthetic vision. We frame the problem in the context of blindness and the clinical data of what implantees see when stimulated using implantable prosthetic vision.

Blindness and low vision are key causes of disability and health burden, the estimated cost of global vision loss is almost US$3 trillion [1]. Therapies such as implantable prosthetic vision aim to provide a benefit to patients by enabling key abilities that otherwise form barriers to independent living. From focus groups, and other research (e.g., [2]), these include face and expression recognition, reading (including packaging, signs and numbers), and orientation and mobility.

A number of trials have been conducted of electrical stimulation of the human visual system. One of the earliest reports is from D'Arsonval in 1896 with an electric coil held near the head. Research in implantable visual prosthetic devices began in earnest in the 1960's when Brindley and Lewin implanted a cortical stimulating device in a human subject [3]. Implants have been trialled in human subjects also in the optic nerve [4] and several positions in the retina, specifically, suprachoroidal, (e.g., Bionic Vision Australia, clinicaltrials.gov/show/NCT01603576), subretinal (e.g., [5]), and epiretinal (e.g., [6]). Other prosthetic vision devices that do not require implantation have been proposed, particularly using sensory substitution [7], including object sonification [8] and tactile and tongue displays. The vision processing approaches described here can also be relevant to a broader category of assistive vision devices.

A phosphene is a sensation of light produced by direct stimulation other than by light[1]. This is electrical stimulation in implanted prosthetic vision. We briefly summarise reports of implantable prosthetic vision, particularly from the multi-centre trials, see [9] for a review. Phosphenes are often small in size and appear in a consistent retinal position. With varying parameters of stimulation, implantees report being able to differentiate apparent levels of intensity. [10] reports patient drawings showing mostly round, and occasionally elongated shapes, with patients able to resolve up to 5-7 levels of noticeable difference. [11] shows different shapes of phosphene, which change shape with stimulus parameters. [12] reports a spatial map of phosphenes that results from an implanted 16 electrode epiretinal array, showing single phosphenes resulting from most electrodes, at locations that largely follow array position (with some distortion and re-ordering), and up to 10 separate levels of noticeable difference. There are also limitations to speed of stimulation, such as fading where sustained stimulation fades out. Zrenner [5] reports that most subjects had an update frequency set to 5Hz, with one as low as 1Hz.

Acuity tests measure the spatial resolution of vision. [6] reports grating acuity results with 28 participants, implanted with a 6x10 electrode array that projects to 20 degrees in the diagonal on the retina, with a matching angle of input from the camera. Seven participants reported acuity above screening value, and the best reported acuity of 1.8 logMAR, approximately 1 cycle per degree. This is greater than would be predicted by sampling theory for a single image, and matches results reported for simulated prosthetic vision adjusted for phosphene density (e.g., [13, 14]). In [5], a diagonal 1500 electrode subretinal array projects to 15 degrees retinally, with a matching angle from the camera. The best grating

NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications, and the Digital Economy, and the Australian Research Council (ARC) through the ICT Centre of Excellence Program. This research was also supported in part by ARC through its Special Research Initiative (SRI) in Bionic Vision Science and Technology grant to Bionic Vision Australia (BVA).

[1] Oxford Dictionary

acuity result reported was 3.3 cycles per degree, though the authors treat this result with caution.

We may summarise the implications of implantee perception reports as follows. The best results show an ability to resolve the world at approximately one phosphene per electrode. There is much variation in the appearance of phosphenes, the simplest form reported is approximately circular, which [9] uses as a standard model for simulated prosthetic vision. Phosphenes lead to at most around 10 levels of noticeable difference. The phosphenes are approximately aligned with the retinal position of the stimulating electrode.

2. CHALLENGES FOR PROSTHETIC VISION

[15] defines vision processing for prosthetic vision as a downsampling function, particularly in space and dynamic range. Each phosphene represents a finite area of visual space. The role of vision processing is to preserve information that enables tasks that are critical to restoration of independent living and quality of life of the individual. Generally, information may be lost due to an object being smaller than the corresponding visual field. In many controlled experiments in the literature, the scale of the visual task is controlled by using large visual targets at a distance, or having small distances to the workspace so that objects are clear in the phosphene image. This is not generally possible in real-world tasks. Further loss can occur in the reduction of dynamic range. Most trials in the literature are conducted in so-called high contrast environments. These are largely black and white, and thus likely to be quantised to different levels in a naive dynamic range downsample. Again, most real-world environments contain elements that are not high contrast. The challenge here for vision processing is to perform adaptive quantisation so that critical contrast is preserved. For example, keeping a low contrast trip hazard visible in the perceived phosphene image by ensuring the corresponding phosphene has distinct contrast from its neighbours.

The dynamic range problem presented by prosthetic vision is closely related to tone mapping, where high dynamic range images are reduced for visualisation on standard visual displays. However, rather than reducing from large numbers of input image levels to a few hundred, in prosthetic vision it must be reduced to ten or less. Tone mapping approaches such as [16] ensure large local contrast in the input image is preserved in the output image, but directly, these will only help if key objects are well contrasted in the input image. Otherwise, this requires a scene understanding approach to identify critical objects, and then an approach to ensure contrast between those and their background is preserved. It is also then important that such boundaries are stable in video to prevent flicker in the output display. Histogram equalisation is not adequate as object/background contrast may still be lost.

Key issues may be summarised as:

- the trade-off between field of view and resolution;
- scene understanding to identify and segment semantically important objects to ensure visibility;
- object aware quantisation that maps contrast of important objects to the perceptual experience of the individual; and,
- maintaining stability of the display over time, including preventing flicker in video.

It should be pointed out that limitations of dynamic range and resolution apply to many possible visual prosthetic devices, including sensory substitution devices.

3. PERFORMANCE IN CLINICAL TRIALS

The two multi-centre trials show the most significant functional vision outcome results to date. In Zrenner et al., [5], most participants were successful at tasks including: light perception, light localisation, motion detection, and grating resolution. 5/9 patients reported implant-mediated visual perceptions in daily life. For other tasks, the success numbers were lower, for example, letter reading 4/9. Humayun et al., [6] reports percentage successful on tasks of 28 patients: object localisation (96%), motion discrimination (57%), and discrimination of oriented gratings (23%). For orientation and mobility, overall participants showed statistically significant improvement vs system off on tasks of moving to a black door across a white room, and following a contrasted line.

4. CURRENT PROGRESS IN VISION PROCESSING FOR IMPLANTABLE PROSTHETIC VISION

As the number of phosphenes is low compared to the incoming image, aliasing can be an issue in downsampling. A classical image processing approach is use an antialias filter to remove frequencies above the Nyquist limit. Both Lanczos2 [14] and local averaging [13] have been applied and show improved acuity test results with filtering in simulated prosthetic vision. These are consistent with acuity results in the human implanted trials described above.

Other image processing approaches have been proposed to address downsampling in some way. These include sets of spatio-temporal filters [17], contrast and brightness enhancement, histogram equalisation, edge detection, and morphological operators (e.g., [18, 19]). Papers have also proposed biologically inspired approaches including retinal processing [20], including the silicon retina [21], and space variant processing [22].

Specialised processing approaches have also been developed. [23] propose anisotropic filtering and negative image overlay to give a type of cartoonisation, and apply image expansion to regions labelled by bottom-up saliency. Saliency was proposed for cuing implantees to important objects that are outside the field of view [24], and to highlight ground-based trip-hazards [25]. Obstacle detection and simultaneous localisation and mapping were applied to guide a user on a safe path using vibration motors on the shoulders to supplement implanted prosthetic vision [26]. Obstacle avoidance has also been demonstrated using a fast technique for segmenting isodisparity curves to find a ground-plane and the boundary with objects [27]. Objects and boundaries are then augmented, ensuring boundary contrast is maintained in the quantisation to phosphene levels [28]. Just noticeable difference optimisation can ensure object boundary visibility, while still maintaining efficient use of the limited dynamic range of phosphenes [29]. Using an RGB-D camera and tracking, a system that displays depth and contrast edges in a phosphene-based display that simulates a cortical device was shown by [30]. Also, for navigation, a method that displays the time-to-contact of incoming objects as the intensity of phosphenes has been proposed [31].

Face recognition is a key task, and modest improvements in performance have been shown in simulation with moderate numbers of electrodes [32]. Here some edge enhancement was shown to potentially improve performance. To assist resolution problems, interactive zooming for faces (or other known objects) fixation has been proposed [33]. Vision processing for a non-implantable face detection and recognition system was demonstrated in [34]. Segmentation followed by forcing contrast can ensure objects with low contrast boundaries can be clearly seen in prosthetic vision [35].

Vision processing can also be incorporated into non-implantable visual prosthetic devices. Proposed systems include: classifying fabric type to allow a user to be told the textural appearance of cloth [36]; currency reading [37]; and, tracking of movable objects in the environment [38].

5. EVALUATING PERFORMANCE OF VISION PROCESSING FOR IMPLANTED PROSTHETIC VISION

Little information is available about the performance of vision processing in human implanted studies. The Zrenner et al [5] trials use directly connected photodiodes so vision processing is not currently possible. [6] describes vision processing performed in the 2012 Second Sight trial as a direct downsample of the image corresponding to the field of view, but no details of the downsampling approach used are provided. A number of the approaches from the previous section have been evaluated in human trials using simulated prosthetic vision or some other non-implantable method, including [24, 14, 39, 26, 32]. Most other approaches have been demonstrated in principal.

Simulation is a key approach to evaluation, and will always retain a role in evaluating efficacy and exploring new approaches. Chen et al., [9] developed a software simulation from their model of implanted visual prosthetic psychophysics: round phosphenes (generally Gaussians) that vary with brightness and size with increased intensity resulting from changing stimulation parameters. Lieby et al. [40] contribute fast methods for simulation rendering showing a real-time display for large numbers of phosphenes.

6. THE FUTURE OF VISION PROCESSING FOR IMPLANTABLE PROSTHETIC VISION

Currently, a number of groups are conducting human trials, with products now available. However, significant progress is required so that implantees can use this vision effectively to complete a range of everyday tasks. Although the number of electrodes and phosphenes that implantees see will expand slowly, this is not enough in itself. Vision processing can contribute to addressing the gap by allowing individuals access to key information when the number of phosphenes and levels are limited.

Over the next five to ten years we will see devices emerge that improve performance for tasks that are important to quality of life of the individual. The key problems for the future of vision processing is ensuring the fidelity and information content and stability over time of the visual experience of the implantee. Vision processing must identify objects important to task performance and ensure their visibility. This means understanding and augmenting the scene, however, any model cannot become misleading through being out of coordination with the current image, and so must be used to modulate the image through augmentation rather than replacing the image.

References

[1] Access Economics, "The global economic cost of visual impairment," Tech. Rep., Access Economics, prepared for AMD Alliance International, 2010.

[2] J E Keeffe, K L Francis, C D Luu, N Barnes, E L Lamoureaux, and R H Guymer, "Expectations of a visual prosthesis: perspectives from people with impaired vision," in *ARVO*, May 2010.

[3] G S Brindley and W S Lewin, "The sensations produced by electrical stimulation of the visual cortex," *Journal of Physiology*, vol. 196, no. 2, pp. 479–493, May 1968.

[4] J Delbeke, M Oozeer, and C Veraart, "Position, size and luminosity of phosphenes generated by direct optic nerve stimulation," *Vision Research*, vol. 43, pp. 1091–1102, 2003.

[5] K Stingl, KU Bartz-Schmidt, D Besch, A Braun, A Bruckmann, F Gekeler, U Greppmaier, S Hipp, G Hortdorfer, C Kernstock, A Koitschev, A Kusnyerik, H Sachs, A Schatz, K T Stingl, T Peters B Wilhelm, and E Zrenner, "Artificial vision with wirelessly powered subretinal electronic implant alpha-ims," *Proc. Royal Society of London B: Biological Sciences*, vol. 280, April 2013.

[6] M S Humayun, J D Dorn, L Da Cruz, G Dagnelie, J-A Sachel, P E Stanga, A V Cideciyan, J L Duncan, D Eliott, E Filley, A C Ho, A Santos, A B Safran, A Arditi, L V Del Priore, and

R J Greenberg for the Argus II Study Group, “Interim results from the international trial of second sight's visual prosthesis,” *Ophthalmology*, vol. 119, no. 4, pp. 779–88, April 2012.

[7] P Bach y Rita and S W Kercel, “Sensory substitution and the human -machine interface,” *Trends in Cognitive Sciences*, vol. 7, no. 12, pp. 541–6, Dec. 2003.

[8] F Ribeiro, D Florencio, P A Chou, and Z Zhang, “Auditory augmented reality: Object sonification for the visually impaired,” in *IEEE-MMSP'12*, 2012.

[9] S C Chen, G J Suaning, J W Morley, and N H Lovell, “Simulating prosthetic vision: 1 visual models of phosphenes,” *Vision Research*, vol. 49, no. 12, pp. 1493–1506, 2009, DOI:10.1016/j.visres.2009.02.003.

[10] R Wilke, V-P Gabel, H Sachs, K-U Bartz-Scmidt, F Gekeler, D Besch, P Szurman, A Stett, B Wilhelm, T Peters, A Harscher, U Greppmaier, S Kibbel, H Benav, A Bruckmann, K Stingl, A Kusnyerik, and E Zrenner, “Spatial resolution and perception of patterns mediated by a subretinal 16-electrode array in patients blinded by hereditary retinal distrophyies,” *Investigative Ophthalmology and Visual Science*, vol. 52, no. 8, pp. 5995–6003, June 2011.

[11] D Nanduri, M Humayun, R Greenberg, M J McMahon, and J Weiland, “Retinal prosthesis shape analysis,” in *IEEE EMBC*, 2008, pp. 1785–8.

[12] M S Humayun, J Weiland, G Y Fujii, R Greenberg, R Williamson, J Little, B Mech, V Cimmarusti, G Van Boemel, G Dagnelie, and E deJuan, “Visual perception in a blind subject with a chronic microelectronic retinal prosthesis,” *Vision Research*, vol. 43, pp. 2573–2585, 2003.

[13] S C Chen, L E Hallum, N H Lovell, and G J Suaning, “Visual acuity measurement of prosthetic vision: A virtual-reality simulation study,” *Neural Engineering*, vol. 2, no. 1, pp. S134–45, Feb 2005.

[14] P Lieby, N Barnes, J G Walker, A F Scott, and L Ayton, “Evaluating Lanczos2 image filtering for visual acuity in simulated prosthetic vision,” in *ARVO*, May 2013.

[15] N Barnes, X He, C McCarthy, L Horne, J Kim, A F Scott, and P Lieby, “The role of vision processing in prosthetic vision,” in *IEEE EMBC*, 2012.

[16] S Ferradans, M Beralmio, E Provenzi, and V Caselles, “An analysis of visual adaptation and contrast perception for tone mapping,” *IEEE T-PAMI*, vol. 33, no. 10, pp. 2002–2012, 2011.

[17] R Hornig, T Zehnder, M Veikay-Perel, T Laube, M Feucht, and G Richard, *The IMI Retinal Implant System*, pp. 111–128, Springer, Berlin, 2001.

[18] W Yu, Y Tian, X Chai, and Q Ren, “Image processing strategies dedicated to the optic nerve stimulation,” in *International Symposium on Biophotonics, Nanophotonics, and Metamaterials*, Oct. 2006, pp. 147–150.

[19] W Liu, W Fink, MA Tarbell, and M Sivaprakasam, “Image processing and interface for retinal visual prostheses,” in *ICAS*, 2005, pp. 2927–2930.

[20] A Asher, W A Segal, S A Baccus, L P Yaroslavsky, and D V Palanker, “Image processing for a high-resolution optoelectronic retinal prosthesis,” *IEEE Trans. Biomedical Engineering*, vol. 54, no. 6, pp. 993–1004, Jun. 2007.

[21] K A Zaghloul and K Boahen, “A silicon retina that reproduces signals in the optic nerve,” *Neural Engineering*, vol. 3, no. 4, pp. 257–67, 2006.

[22] M Vurro, G Baselli, F Orabana, and G Sandini, “Simulation and assessment of bioinspired visual processing system for epiretinal prostheses,” in *IEEE EMBC*, 2006, pp. 3278–81.

[23] W Al-Atabany, B McGovern, K Mehran, R B Palmini, and P Degenaar, “A processing platform for optoelectronic/optogenetic retinal prosthesis,” *IEEE Trans. Biomedical Engineering*, vol. 54, no. 6, pp. 1–10, Nov. 2011.

[24] N Parikh, L Itti, and J Weiland, “Saliency-based image processing for retinal prosthesis,” *Neural Engineering*, vol. 7, no. 1, pp. 1–10, Jan 2010.

[25] A Stacey, Y Li, and N Barnes, “A salient information processing system for bionic eye with application to obstacle avoidance,” in *IEEE EMBC*, 2011.

[26] J D Weiland, N Parikh, V Pradeep, and G Medioni, “Smart image processing system for retinal prosthesis,” in *IEEE EMBC*, 2012, pp. 300–303.

[27] C McCarthy and N Barnes, “Surface extraction with isodisparity contours,” in *Proc. ACCV*, Queenstown, NZ, 2010.

[28] C McCarthy, N Barnes, and P Lieby, “Ground surface segmentation for navigation with a low resolution visual prosthesis,” in *IEEE EMBC*, 2011.

[29] Y Li, C McCarthy, and N Barnes, “On just noticeable difference for bionic eye,” in *IEEE EMBC*, 2012.

[30] W L D Lui, D Browne, L Kleeman, and H L Wai, “Transformative reality: Improving bionic vision with robotic sensing,” in *IEEE EMBC*, 2012, pp. 304–307.

[31] C McCarthy and N Barnes, “Time-to-contact maps for navigation with a low resolution visual prosthesis,” in *IEEE EMBC*, 2012.

[32] M H Chang, H S Kim, J H Shin, and K S Park, “Facial identification in very low-resolution images simulating prosthetic vision,” *Neural Engineering*, vol. 9, no. 1, 2012.

[33] X He, J Kim, and N Barnes, “An face-based visual fixation system for prosthetic vision,” in *IEEE EMBC*, 2012.

[34] S Krishna, G Little, J Black, and S Panchanathan, “iCARE interaction assistant: A wearable face recognition systems for individuals with visual impairments,” in *ASSETS'05*, 2005.

[35] L Horne, N Barnes, C McCarthy, and X He, “Image segmentation for enhancing symbol recognition in prosthetic vision,” in *IEEE EMBC*, 2012.

[36] X Yang, Yuan, and Y Tian, “Recognizing clothes patterns for blind people by confidence margin based feature combination,” in *MM'11*, 2011.

[37] X Lu, “A camera phone based currency reader for the visually impaired,” in *ASSETS'08*, 2008.

[38] A Hub, T Hartter, and T Ertl, “Interactive tracking of movable objects for the blind on the basis of environment models and perception-oriented object recognition models,” in *ASSETS'06*, 2006.

[39] C McCarthy, P Lieby, J G Walker, A F Scott, V Botea, and N Barnes, “Low contrast trip hazard avoidance with simulated prosthetic vision,” in *ARVO*, May 2012.

[40] P Lieby, N Barnes, C McCarthy, N Liu, H Dennnett, J G Walker, V Botea, and A F Scott, “Substituting depth for intensity and real-time phosphene rendering: Visual navigation under low vision conditions,” in *IEEE EMBC*, 2011.