

Towards Photosensor Movement-Adaptive Image Analysis in an Electronic Retinal Prosthesis

L.E. Hallum^{1,2}, G.J. Suaning^{3,1}, D.S. Taubman², and N.H. Lovell¹

¹ Graduate School of Biomedical Engineering,

² School of Electrical Engineering and Telecommunications,
University of New South Wales, Sydney, Australia, 2052.

³ School of Engineering,
University of Newcastle, Newcastle, Australia, 2308.

Abstract— By way of extracellular, stimulating electrodes, electronic vision prosthesis aims to render discrete light spots – so-called phosphenes – in the visual field, thereby providing a phosphene image serving as a rudimentary remediation of profound blindness. It is proposed that a digital camera, or some other photosensitive array, captures frames, the frames be analyzed, and phosphenes be actuated accordingly. We present a numerical experiment wherein we observed the phosphene image in response to a set of stimuli for various image analysis schemes. We used the mutual-information function to quantify the efficacy of analysis schemes; the function penalizes a scheme for introducing redundancy to the phosphene image, while accounting for the probability of each stimulus. We demonstrate an effective scheme involving Laplacian of Gaussian ($\nabla^2 G$) kernels geometrically transformed in accordance with phosphene layout. Further, we propose adapting the kernels comprising a scheme in accordance with photosensor movement.

Keywords— retinal prosthesis; artificial vision; spatial filtering; mutual-information function.

I. INTRODUCTION

EXISTING studies seek to quantify the usefulness of prosthetic vision for performing daily tasks – e.g., mobility [1], reading speed [2], face recognition [3], and household activities [4]. There is scarce treatment in the literature, however, of image analysis as applied to prosthetic vision, for integration with a prosthesis.

One such treatment is previous work by the authors wherein we assessed the fixation and pursuit capabilities of subjects afforded simulated prosthetic vision [5]. In this previous experiment, subjects were required to track a small, high-contrast target, at times stationary and at times moving, with the central phosphene of a phosphene array, the movement of which subjects effected via a joystick. Throughout, each subject’s deviation signal was recorded: a discrete time-series describing the displacement of the central phosphene from the target. Subjects performed approximately three hours’ tracking with three image analysis schemes. An important observation was made: subjects used the non-trivial analysis scheme (involving overlapping Gaussian kernels chosen *ad hoc*) to their advantage over the trivial scheme of the existing literature: regional averaging.

The present paper concerns the informed design of an image analysis scheme for a retinal prosthesis. In the same way that speech processing in the cochlear implant assists recognition of consonants, vowels, monosyllabic words, and sentences [6], a well designed image analysis scheme should improve the artificial vision afforded prosthetic retinal implantees.

II. METHOD

Fig. 1 depicts the numerical set-up. The impulse is shifted through all locations in the numerical region, R (lower layer). For each impulse location, the analysis scheme, Q (middle layer), is applied to R , determining the levels of activation of phosphenes in the phosphene image (upper layer).

A scheme comprises seven spatial kernels, arranged, like phosphenes, in an hexagonal mosaic, each of which analyses a subregion of R and drives a phosphene; we employ seven phosphenes because previous work shows that, in the main, fixation and pursuit is achieved with the use of these central phosphenes. It is convenient to think, therefore, of the biological analogy: a subregion as a phosphene’s receptive field, and the kernel as the receptive field profile.

Phosphene output was quantized to 32 levels in accordance with our neurostimulator design, central to which is a five-bit DAC [7]. (It is worth noting that, by varying the pulse duration of stimulation, a greater number of percepts may be achievable. For a lack of *in vivo* data in the literature, our selection of 32 is somewhat speculative.) We define phosphene pitch (PP) as the shortest distance between any two phosphenes given the hexagonal mosaic used. We ran the experiment for various schemes, Q : regional averaging (RA) schemes, Gaussian (G) schemes, and Laplacian of Gaussian ($\nabla^2 G$) schemes, as discussed subsequently.

We used the mutual-information function to assess the effectiveness of each analysis scheme. In the present context, the input, \mathbf{X} , describes the location of the impulse; the output is a vector of dimensionality seven, \mathbf{Y} , describing the activations of phosphenes comprising the phosphene image. Thus, the

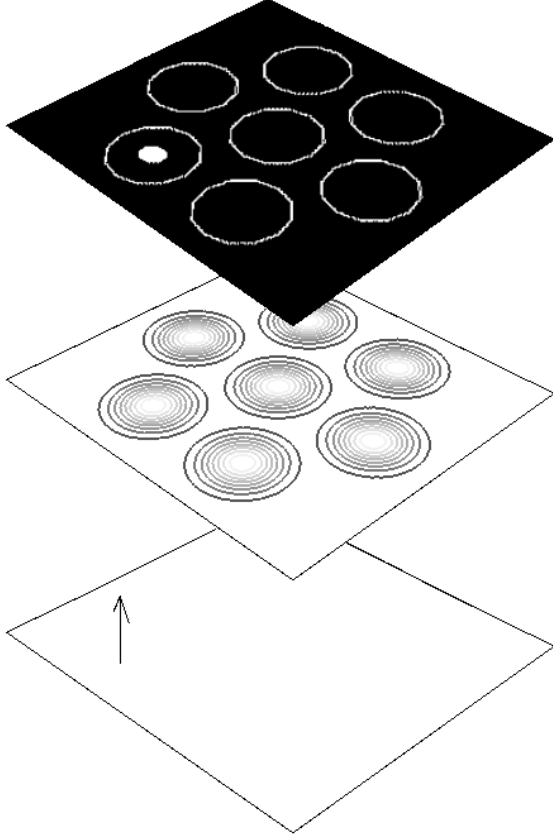


Fig. 1. The numerical set-up. The impulsive stimulus (bottom layer) is analyzed by filtering scheme, Q (middle layer), which in turn excites the phosphene image (upper layer). Note that the scheme shown involves non-overlapping Gaussian kernels.

mutual-information function,

$$I(\mathbf{X}; \mathbf{Y}) = H(\mathbf{X}) - H(\mathbf{X}|\mathbf{Y}) \quad \text{bits/symbol}, \quad (1)$$

is equal to the entropy of the input,

$$H(\mathbf{X}) = - \sum_{j=0}^{J-1} p_j \log_2 p_j \quad \text{bits/symbol},$$

less the conditional entropy,

$$H(\mathbf{X}|\mathbf{Y}) = - \sum_{k=0}^{K-1} q_k \sum_{j=0}^{J-1} P_{j|k} \log_2 P_{j|k} \quad \text{bits/symbol},$$

where $P_{j|k}$ is the backwards transition matrix of the analysis scheme, Q , and $q_k = \sum_{j=0}^{J-1} p_j Q_{k|j}$.

That is, the function measures the degree to which seeing the phosphene image reduces the uncertainty as to the location of the impulse.

Central to the mutual-information function is the probability density function associated with the input vector, \mathbf{X} . This density function, in effect, ascribes a probability, p_j , to each impulse based on its location. In analysing data from the previous experiment described briefly in the Introduction [5], we noticed that the first-order statistics of the random process comprising subjects' deviation signals are well modelled by a bivariate normal [8]. That is, subjects tended to deviate about the target, be it stationary or moving, so as to describe a bivariate normal. Thus, we can ascribe a probability, p_j , to each impulse based on its horizontal, x_1 , and vertical, x_2 , position in the numerical region, where $(x_1, x_2) = (0, 0)$ is the region's centre, as follows:

$$P(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} \exp\left[-\frac{z}{2(1-\rho^2)}\right], \quad (2)$$

where,

$$z \equiv \frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1 x_2}{\sigma_1 \sigma_2} + \frac{x_2^2}{\sigma_2^2},$$

ρ is the correlation of x_1 and x_2 , and σ_1^2 and σ_2^2 are the variances of the normal in the horizontal and vertical directions respectively.

Deviation signals from [5] were modelled using an implementation of the non-linear least-squares Levenberg-Marquardt algorithm [9]. In short, for fixation, we found $\rho \approx 0.0$, $\sigma_1 = 0.14$ PP and $\sigma_2 = 0.13$ PP; for pursuit, $\rho \approx 0.0$, $\sigma_1 = 0.26$ PP and $\sigma_2 = 0.27$ PP.

An analysis scheme will affect the breadth of scanning, $\sigma_{1,2}$; therefore, have we constructed a circular argument? It is important to note that, while the first-order statistics of the deviation process varied for fixation and pursuit, and for different schemes, the normal nature of the processes remained. For example, mean pursuit deviation for regional averaging was greater than that of Gaussian kernels, but both are well modelled by bivariate normals (just of different variances). Subsequent paragraphs assess filtering schemes for a range of breadth of scanning so as to account for the effect of the scheme on scanning.

By this means – shifting the impulse through the numerical region and, in light of the probability density function of the impulse's location, using the mutual-information function to

assess information content in the phosphene image – we have assessed different analysis schemes, Q .

III. RESULTS

Fig. 2 demonstrates the relative efficacy of three filtering schemes. The first, RA , is comprised of hexagonal kernels, each of which averages the intensity of its receptive field. For RA , the kernel extent (x -axis) refers to the height of the hexagon in question; therefore, for extent = 1.0 PP, the numerical region is effectively tiled by hexagonal kernels. The second, G , is comprised of Gaussian kernels. The third, LoG , is comprised of $\nabla^2 G$ kernels. For G and $\nabla^2 G$, kernel extent refers to the radial standard deviation of the kernel in question.

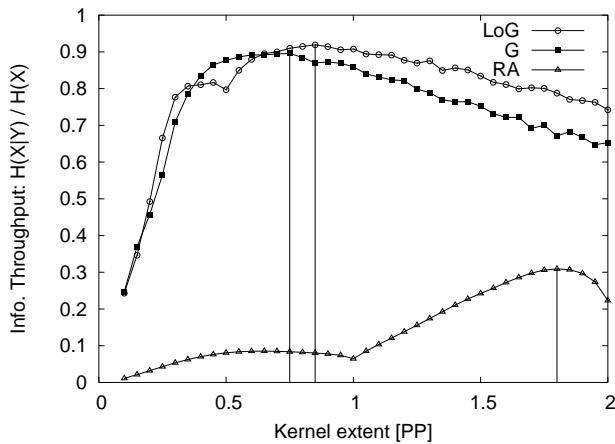


Fig. 2. Relative efficacy of three analysis schemes: regional averaging (RA), Gaussian kernels (G), and $\nabla^2 G$ kernels (LoG) for $\sigma_1 = \sigma_2 = 0.25$ PP. Vertical lines indicate global maximum for each curve. Thus, a scheme of $\nabla^2 G$ kernels with standard deviation = 0.85 PP is best of those shown.

From the figure we can discern that, for breadth of scanning characterized by $\sigma_1 = \sigma_2 = 0.25$ PP, a scheme comprising $\nabla^2 G$ kernels of standard deviation 0.85 PP will convey most information to an observer of the phosphene image.

As shown in the upper layer of Fig. 1, for the purposes of this experiment, we have arranged phosphenes so as to form an hexagonal mosaic. (We have explained elsewhere the likely benefit of manufacturing implantable electrode arrays to form hexagonal mosaics [10].) This gives rise to the question, Can the geometry of the phosphene layout be incorporated into the geometry of the kernels comprising the analysis scheme so as to increase the information put through to the phosphene image? Fig. 3 shows a $\nabla^2 G$ kernel (left) and its hexagonalized equivalent (right). Intuitively, hexagonalized kernels make for a “better fit” when arranged in an hexagonal mosaic. For schemes comprised of hexagonal $\nabla^2 G$ kernels we re-ran the

numerical experiment, taking note, at each breadth of scanning, of the information put through by the optimal G scheme, and the optimal $\nabla^2 G$ scheme. Fig. 4 demonstrates the efficacy of the hexagonal $\nabla^2 G$ kernel: for scanning characteristic of both fixation and pursuit (that is, for breadth of scanning ranging from 0.05 to 0.5 PP), an hexagonal $\nabla^2 G$ scheme will convey more information to the observer of the phosphene image.

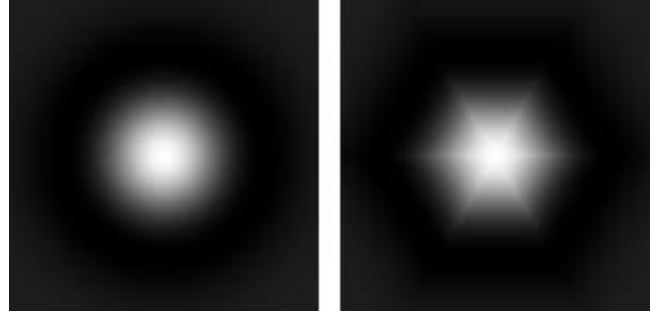


Fig. 3. The circularly symmetric $\nabla^2 G$ kernel (left) and its hexagonally symmetric equivalent, formed by geometric transformation (right).

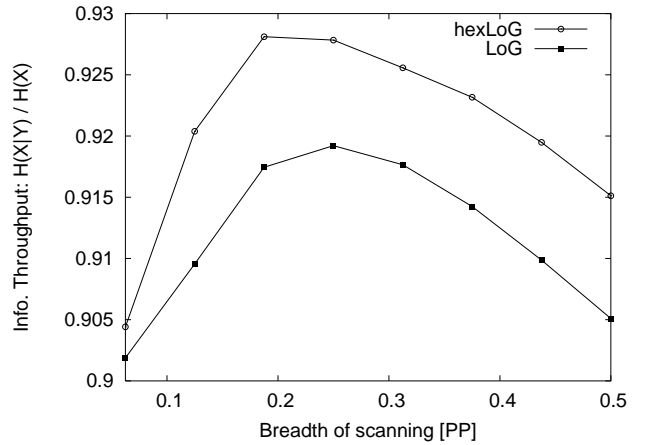


Fig. 4. The efficacy of the hexagonally symmetric $\nabla^2 G$ kernel as compared with its circularly symmetric equivalent. An analysis scheme comprising hexagonally symmetric $\nabla^2 G$ kernels puts through more information to the observer of the phosphene image.

IV. DISCUSSION

Using the mutual-information function as a measure, we have assessed the effectiveness of different image analysis schemes for use in a retinal prosthesis. We demonstrated that schemes comprising Gaussian kernels are superior to regional averaging in that they convey more information to the observer of the phosphene image. Thus we have provided

an information theoretic reason for our previous empirical result as briefly described in the Introduction [5]. Superior to both Gaussian and regional averaging schemes, however, is a scheme comprising Laplacian of Gaussian ($\nabla^2 G$) kernels (s.d. = 0.85 PP). Further, we have shown that by geometrically transforming kernels in accordance with the geometry of the phosphene layout (here an hexagonal mosaic) still more information is put through to the phosphene image. The benefit of geometric transformation persists for scanning characteristic of fixation and pursuit (indeed, the benefit persists for uniform distribution of the impulse, $P(x_1, x_2)$).

The hexagonalized $\nabla^2 G$ kernels made for modest mutual-information gains, but the technique of geometric transformation will be important where irregular phosphene layouts are concerned, such as those reported by Humayun *et al.* as part of an early clinical trial [11]. Our present results imply the geometric transformation of kernels on Voronoi polygons formed about phosphene centres (a demonstration of which is currently in preparation at the authors' laboratory).

The $\nabla^2 G$ kernel is of interest in the present context; it has attracted interest in biological and image processing fields alike. Initially proposed by Marr and Hildreth in the context of computational visual modelling [12], its being dually concentrated in spatial and frequency domains allows for better localization of more features. Further, its circular symmetry means no information is lost due to orientation.

A useful by-product of this work is as follows: by varying the resolution of the numerical region (bottom layer as per Fig. 1), and the number of levels to which phosphene output is quantized, we can quantify, respectively, the effect of photosensor resolution – be it a digital camera, as proposed by the authors [13], or an implanted photodiode array [14] – on the resolution of the phosphene image, and the effect of the resolution of the implanted neurostimulator's DAC. For a numerical region of decreasing resolution, and for the quantization of phosphene output to fewer levels (say, eight as opposed to 32), the information put through to the phosphene image is decreased.

How is the present work a step towards photosensor movement-adaptive image processing? The set-up presented here can incorporate dynamical aspects of the deviation signal, \mathbf{X} , which correlate to photosensor array movement. This addition will likely show how orientation- and scale-specific kernels can be incorporated with the present findings. The use of such kernels would need to be adaptive, otherwise information from the “other” orientations and scales would be omitted from the phosphene image. Said use in the adaptive sense, inasmuch as a single phosphene, when observed over time, would convey information pertinent to numerous scales and orientations, would mimic the dyadic use of Gabor filters,

for example, for texture segmentation in machines [15].

REFERENCES

- [1] K. Cha, K.W. Horch, and R.A. Normann, “Mobility performance with a pixelized vision system,” *Vision Res*, vol. 32, pp. 1367–1372, 1992.
- [2] K. Cha, K.W. Horch, R.A. Normann, and D.K. Boman, “Reading speed with a pixelized vision system,” *J Op Soc Am A*, vol. 9, pp. 673–677, 1992.
- [3] Jr R.W. Thompson, G.D. Barnett, M.S. Humayun, and G. Dagnelie, “Facial recognition using simulated prosthetic vision,” *Invest Ophthalmol Vis Sci*, vol. 44, pp. 5035–5042, 2003.
- [4] M.S. Humayun, “Intraocular retinal prosthesis,” *Tr Am Ophth Soc*, vol. 99, pp. 271–300, 2001.
- [5] L.E. Hallum, G.J. Suaning, and N.H. Lovell, “Effecting improved prosthetic visual acuity via image processing in a fixation-pursuit task,” *Invest Ophthalmol Vis Sci*, ARVO abstract 4175, 2004.
- [6] L.A. Whitford, P.M. Seligman, C.E. Everingham, T. Antognelli, M.C. Skok, R.D. Hollow, K.L. Plant, E.S. Gerin, S.J. Staller, and H.J. McDermott, “Evaluation of the Nucleus Spectra 22 processor and new speech processing strategy (SPEAK) in postlingually deafened adults,” *Acta Otolaryngol*, vol. 115, pp. 629–637, 1995.
- [7] G.J. Suaning and N.H. Lovell, “CMOS neurostimulation ASIC with 100 channels, scaleable output, and bi-directional radio-frequency telemetry,” *IEEE Trans Biomed Eng*, vol. 48, pp. 248–260, 2001.
- [8] L.E. Hallum, G.J. Suaning, and N.H. Lovell, “Simulated prosthetic visual fixation, saccade, and smooth pursuit; and the use of non-trivial image processing to effect improved prosthetic vision,” Unpublished data.
- [9] M. Galassi, J. Davies, J. Theiler, B. Gough, J. Jungman, M. Booth, and F. Rossi, “GNU Scientific Library - Routines for numerical computing (version 1.4),” 2003, Computer software.
- [10] L.E. Hallum, D.S. Taubman, G.J. Suaning, J.W. Morley, and N.H. Lovell, “A filtering approach to artificial vision: a phosphene visual tracking task,” in *Proc. World Congress of Biomedical Engineering and Medical Physics*, Sydney, Australia, 2003.
- [11] M.S. Humayun, J.D. Weiland, G.Y. Fujii, R. Greenberg, R. Williamson, J. Little, B. Mech, V. Cimarusti, G. Van Boemel, G. Dagnelie, and E. de Juan, “Visual perception in a blind subject with a chronic microelectronic retinal prosthesis,” *Vision Res*, vol. 43, pp. 2573–2581, 2003.
- [12] D. Marr and E. Hildreth, “Theory of edge detection,” *Proc R Soc Lond B Biol Sci*, vol. 207, pp. 187–217, 1980.
- [13] G.J. Suaning, L.E. Hallum, S.C. Chen, P.J. Preston, and N.H. Lovell, “Phosphene vision: Development of a portable visual prosthesis system for the blind,” in *Proc. 25th Annual International Conference of the IEEE/EMBS*, Cancun, Mexico, 2003.
- [14] H.N. Schwahn, F. Gekeler, K. Kohler, K. Kobuch, H.G. Sachs, F. Schulmeyer, W. Jakob, V.P. Gabel, and E. Zrenner, “Studies on the feasibility of a subretinal visual prosthesis: data from yucatan micropig and rabbit,” *Graefes Arch Clin Exp Ophthalmol*, vol. 239, pp. 961–967, 2001.
- [15] A.K. Jain and F. Farrokhnia, “Unsupervised texture segmentation using Gabor filters,” *Pattern Recogn*, vol. 24, pp. 1167–1186, 1991.