

Psychophysics of Prosthetic Vision: II. Stochastic Sampling, the Phosphene Image, and Noise

Luke E. Hallum, Spencer C. Chen, Shaun L. Cloherty, and Nigel H. Lovell

Abstract—Stimulation of the diseased retina via an intraocular electrode array is a proposed means of restoring some vision to the profoundly blind. A prosthetic device to this end would involve post-implantation calibration (analogous to cochlear implant fitting), wherein the subject indicates those discrete positions in the visual field where luminous percepts are elicited. This procedure would be a source of noise, because the indicated positions would only approximate the actual positions in the visual field. Put differently, the procedure introduces sampling jitter, and would therefore affect clinical outcomes such as mobility and reading speeds. The nature of this noise is the concern of the present paper; we derive an expression for the noise power spectrum as it relates to the statistical nature of the sampling jitter. We show that, generally, jitter has greater effect on higher spatial-frequencies, that is, those areas of the implantee's visual perception that represent fine detail are more prone to noise. More specifically, the noise spectrum depends on the characteristic function of the random variable describing the sampling jitter. Our results signal the need for experimental work that characterizes sampling jitter in implantees, plus the need for simulations that allow a better understanding of perception and the noisy phosphene image.

I. INTRODUCTION

THE microelectronic retinal prosthesis (for review see [1]) proposes, by way of an intraocular array of stimulating electrodes, to render luminous spots (so-called phosphenes which together comprise the phosphene image) in fixed locations of the visual field of the profoundly blind implantee. By way of analogy, in the cochlear implant, the implanted subject undergoes post-surgical, psychophysical calibration sessions wherein perceptual thresholds and maximum comfortable loudness are determined at each electrode. We envision a similar procedure for retinal implantees wherein, via psychophysical means, the locations in the visual field occupied by phosphenes are indicated. These locations inform the image processing that lies functionally intermediate to the intraocular electrode array and the digital camera (worn on the body) that acquires real-world scenes; specifically, the locations determine where the real-world scenes are sampled prior to the corresponding activations of electrodes. This calibration procedure would be necessarily inaccurate. That is to say, the phosphene locations as indicated by the subject only approximate the actual locations of phosphenes in the visual field. This introduces noise to

the phosphene image that is more pronounced where high-spatial-frequency information (fine detail) in the real-world scene is concerned. This concept is depicted in Fig. 1. Note that, in the figure, phosphenes are arranged in an orderly fashion in the visual field (of the phosphene image observer). In practice we would expect irregularity in the phosphene layout due to the electrotonicity of the retina (e.g., see [2], [3]).

In the following section, we derive an expression (in one dimension) for the power contained in this noise as a function of spatial-frequency of real-world scenes, and the relationship between the noise and the statistics of the sampling jitter. In the Discussion, we take up the potential clinical significance of this result and signal the need for work concerning perception and the phosphene image in the presence of noise.

II. STATISTICAL ANALYSIS OF SAMPLING JITTER

We have formulated our approach as follows. Firstly, consider a zero-phase sinusoid with frequency u_a on the range $[0, 1]$, that is, $(1 + \cos(2\pi u_a x))/2$, sampled at $x = 0$. If the sample point is subject to a jitter, χ , then the resulting amplitude error is given by $\epsilon = (1 + \cos(2\pi u_a x))/2 - (1 + \cos(2\pi u_a (x + \chi)))/2$ evaluated at $x = 0$. This expression for amplitude error may be generalised as,

$$\epsilon(x, u; A) = \frac{1}{2} \cos(2\pi A) - \frac{1}{2} \cos(2\pi u(x + A/u)), \quad (1)$$

where A , $0 \leq A \leq 1$, is a parameter that, for the meantime, takes the value $A = 0$, but will later be used to account for shifts, over a single period, of the original zero-phase sinusoid. Therefore, the power contained in the error is related to

$$\begin{aligned} \epsilon^2(x, u; A) &= \frac{1}{4} \cos^2(2\pi A) \\ &\quad - \frac{1}{2} \cos(2\pi A) \cos(2\pi u(x + A/u)) \\ &\quad + \frac{1}{4} \cos^2(2\pi u(x + A/u)). \end{aligned}$$

The statistical average over x (see [5, p. 12]), then, of the power contained in the error is given by

$$\begin{aligned} E(\epsilon^2(x, u; A)) &= \frac{1}{4} \cos^2(2\pi A) \int_{-\infty}^{\infty} w(x) dx \\ &\quad - \frac{1}{2} \cos(2\pi A) \int_{-\infty}^{\infty} w(x) \cos(2\pi u(x + A/u)) dx \\ &\quad + \frac{1}{4} \int_{-\infty}^{\infty} w(x) \cos^2(2\pi u(x + A/u)) dx, \quad (2) \end{aligned}$$

L. E. Hallum, S. C. Chen and S. L. Cloherty are with the Graduate School of Biomedical Engineering, University of New South Wales, Sydney, NSW, Australia, 2052.

N. H. Lovell is with the Graduate School of Biomedical Engineering, University of New South Wales, Sydney, NSW, Australia, 2052, and National Information and Communications Technology Australia (NICTA), Eveleigh, NSW, Australia, 1430.

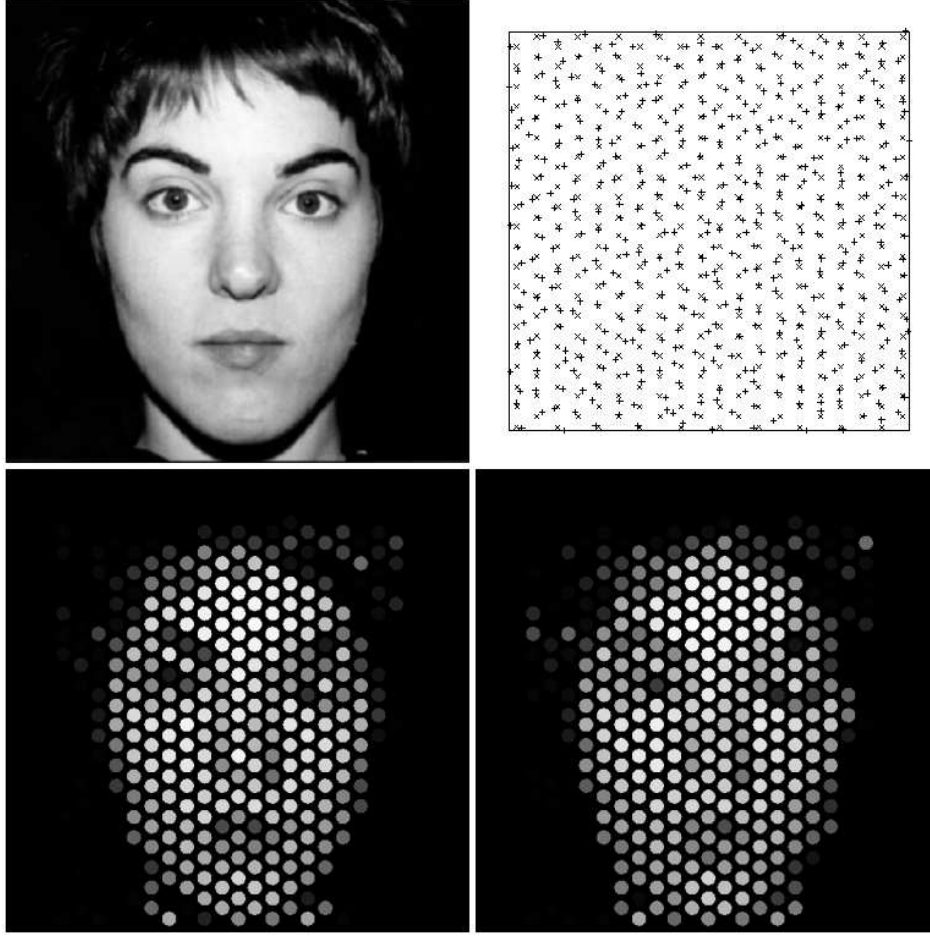


Fig. 1. The effect of sampling jitter on a phosphene image of a face. The face (top, left; image source [4]) is sampled by two mosaics, shown superimposed (top, right); one mosaic is ordered (\times 's) and the other is subject to a small amount of uniform jitter ($+$'s). The samples from each mosaic are then used to render a phosphene image (ordered sampling: bottom, left; jittered sampling: bottom right); note that locations of phosphenes, which are congruent with the ordered sampling mosaic, do not differ between the two phosphene images. The bottom, right phosphene image is therefore noisy. This noise is most pronounced for fine detail, e.g., the eyes and eyebrows, and the hairline.

where $E(\cdot)$ denotes the expectation operator, and $w(x)$ is the first-order probability density function (p.d.f.) that governs the sampling jitter. Of course, the first term in Eq. (2) simplifies to $\cos^2(2\pi A)/4$ since the integral equates to 1. The second term relates to the Fourier transform (FT) of $w(x)$, that is, the characteristic function (c.f.) of the random variable (r.v.) jitter as follows (see [5, p. 13]):

$$\begin{aligned}
 & \int_{-\infty}^{\infty} w(x) \cos(2\pi u(x + A/u)) dx \\
 &= \int_{-\infty}^{\infty} w(x' - A/u) \cos(2\pi ux') dx' \\
 &= \Re\{\mathcal{F}\{w(x' - A/u)\}\} \\
 &= \hat{w}(u) \cos(2\pi A) \tag{3}
 \end{aligned}$$

where we have used a change of variable, $x' = x + A/u$, \Re denotes 'the real part of', \mathcal{F} denotes 'the Fourier transform of', and the circumflex notation denotes a Fourier pair, $w \leftrightarrow \hat{w}$. These equations involve recognizing that, firstly, $\int_{-\infty}^{\infty} f(x) \cos(2\pi ux) dx$ equates to the real part of the FT of $f(x)$, and that, secondly, shifting w affects the phase of the

FT, \hat{w} (see [6, pp. 111–113]).

The third term in Eq. (2) relates to the constricted c.f. of r.v. jitter like so:

$$\begin{aligned}
 & \int_{-\infty}^{\infty} w(x) \cos^2(2\pi u(x + A/u)) dx \\
 &= \int_{-\infty}^{\infty} w(x' - A/u) \cos^2(2\pi ux') dx' \\
 &= \frac{1}{2} \int_{-\infty}^{\infty} w(x' - A/u) (\cos(4\pi ux') + 1) dx' \\
 &= \frac{1}{2} \int_{-\infty}^{\infty} w(x' - A/u) \cos(4\pi ux') dx' + \frac{1}{2} \\
 &= \frac{1}{2} \int_{-\infty}^{\infty} w(x''/2 - A/u) \cos(2\pi ux'') dx'' + \frac{1}{2} \\
 &= \frac{1}{2} \Re\{\mathcal{F}\{w(x''/2 - A/u)\}\} + \frac{1}{2} \\
 &= \frac{1}{2} \hat{w}(2u) \cos(2\pi A) + \frac{1}{2}, \tag{4}
 \end{aligned}$$

where the approach is much the same as in Eq. (3), and a second change of variable was used: $x' = x''/2$. Then,

Eq. (2) simplifies as follows,

$$E(\epsilon^2(x, u; A)) = \left(\frac{1}{4} - \frac{1}{2} \hat{w}(u) \right) \cos^2(2\pi A) + \frac{1}{8} \hat{w}(2u) \cos(2\pi A) + \frac{1}{8}$$

which, so as to account for all possible (and equally probable) phases of what was originally our zero-phase cosinusoid, we integrate on the interval $A = [0, 1]$ to get

$$\begin{aligned} E(\epsilon^2(x, u)) &= \int_0^1 \left\{ \frac{1}{4} (1 - 2\hat{w}(u)) \cos^2(2\pi A) + \frac{1}{8} \hat{w}(2u) \cos(2\pi A) + \frac{1}{8} \right\} dA \\ &= \frac{1}{4} (1 - \hat{w}(u)). \end{aligned} \quad (5)$$

This equation (Eq. (5)) expresses the power spectrum of the noise. See Fig. 2. One critical observation to be made is the effect of the first-order p.d.f. that governs the sampling jitter, $w(x)$. The FT of $w(x)$, that is, the c.f. of the r.v. jitter, determines the rate of growth of the noise power. With the reciprocal relationship of w and \hat{w} in mind, as the p.d.f. $w(x)$ narrows, noise power grows more slowly with frequency. In the limiting case, when sample positions are determined (subject to no jitter), then $w(x) = \delta(x) \leftrightarrow 1$, where $\delta(x)$ denotes the Dirac delta function, and so $E(\epsilon^2(x, u)) = (1 - 1)/4 = 0$. That is to say, as expected, in the absence of jitter, no noise is introduced to the phosphene image. At the other extreme, when jitter is very broad, then all frequencies are noised except a d.c. offset.

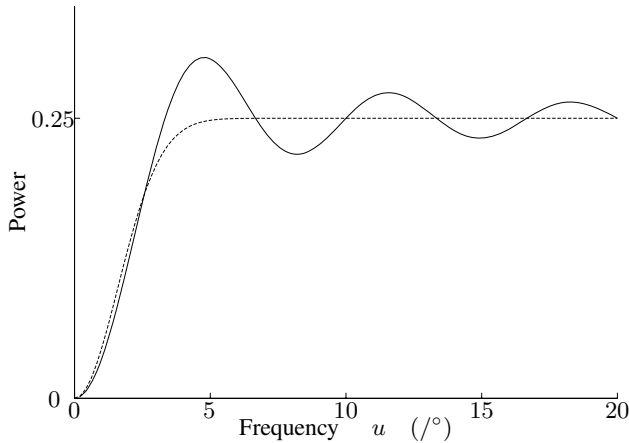


Fig. 2. Growth of noise power with frequency for two classes of sampling jitter: uniform jitter on $[-0.15^\circ, 0.15^\circ]$ (solid line) and normal jitter with standard deviation $\sigma = 0.1^\circ$ (dashed line).

A numerical simulation, using R [7], was undertaken. Analytical and numerical results agreed. At each frequency of a cosine wave on the range $[0, 1]$, and for purely random phase shifts thereof, we computed the mean ($n = 2000$) of the square of the difference of the cosine sampled at $x = 0$, and the same cosine sampled at $x = \chi$. In the plot shown in Fig. 3, χ was uniform on the interval $[-0.15^\circ, 0.15^\circ]$.

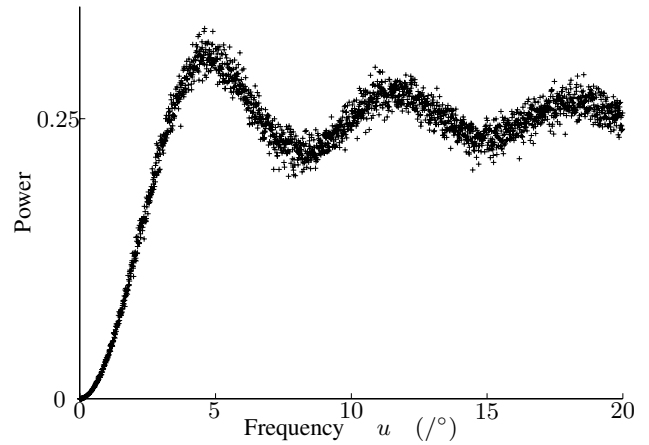


Fig. 3. Numerical simulation results. See text for details. Note the agreement with Fig. 2.

III. DISCUSSION

There is a growing body of literature concerning the phosphene image and perception (e.g., [8], [9], [10]). These studies have involved the recognition of (phosphened) faces, reading speeds of (phosphened) text, and the acuity that the phosphene image affords observers. There is no study, however, that addresses the robustness of perception when faced with a noisy phosphene image. This work is underway in the present authors' lab, and is analogous to a study by Friesen et al. [11], which demonstrated declining speech recognition in both cochlear implant listeners and normally hearing listeners (afforded acoustic simulations) as the signal-to-noise ratio decreased from 15 dB to 0 dB. Similar trends are expected concerning perception of the phosphene image. A body of such data could specify the required accuracy of post-implantation visual calibration with a mind to keeping the adverse perceptual effects of sampling jitter to a minimum.

The foregoing, of course, assumes that the phosphene image observer does not, over the course of practice, learn to compensate completely for sampling jitter. To rephrase, the observer may develop an internal model that effectively denoises the phosphene image. This is also an area of interest to the present authors.

REFERENCES

- [1] E. Zrenner, "Will retinal implants restore vision?" *Science*, vol. 295, pp. 1022–1025, 2002.
- [2] J. Rizzo III, J. Wyatt, J. Loewenstein, S. Kelly, and D. Shire, "Perceptual efficacy of electrical stimulation of human retina with a microelectrode array during short-term surgical trials," *Investigative Ophthalmology & Visual Science*, vol. 44, pp. 5362–5369, 2003.
- [3] M. Humayun, J. Weiland, G. Fujii, R. Greenberg, R. Williamson, and J. L. et al., "Visual perception in a blind subject with a chronic microelectrode retinal prosthesis," *Vision Research*, vol. 43, pp. 2573–2581, 2003.
- [4] P. Hancock, "Psychological Image Collection at Stirling," WWW: <http://pics.psych.stir.ac.uk>, accessed 1 July, 2003.
- [5] D. Middleton, *An Introduction to Statistical Communication Theory*. The Maple Press Company, York, PA, 1960.
- [6] R. Bracewell, *The Fourier Transform and its Applications*. McGraw-Hill, Singapore, 2000.

- [7] R Development Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2004, ISBN 3-900051-00-3. [Online]. Available: <http://www.R-project.org>
- [8] R. Thompson Jr., G. Barnett, M. Humayun, and G. Dagnelie, "Facial recognition using simulated prosthetic vision," *Investigative Ophthalmology & Visual Science*, vol. 44, pp. 5035–5042, 2003.
- [9] A. P. Fornos, J. Sommerhalder, B. Rappaz, A. Safran, and M. Pelizzone, "Simulation of artificial vision, III: Do the spatial or temporal characteristics of stimulus pixelization really matter?" *Investigative Ophthalmology & Visual Science*, vol. 46, pp. 3906–3912, 2005.
- [10] S. Chen, L. Hallum, N. Lovell, and G. Suaning, "Visual acuity measurement of prosthetic vision: a virtual-reality study," *Journal of Neural Engineering*, vol. 2, pp. S135–S145, 2005.
- [11] L. Friesen, R. Shannon, D. Baskent, and X. Wang, "Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants," *Journal of the Acoustical Society of America*, vol. 110, pp. 1150–1163, 2001.