

Psychophysics of Prosthetic Vision: III. Stochastic Rendering, the Phosphene Image, and Perception

Luke E. Hallum, Shaun L. Cloherty, David S. Taubman, Gregg J. Suaning, and Nigel H. Lovell

Abstract—This paper examines the rendering of luminous spots ("phosphenes") in the visual field, and their stochastic positioning as a means of anti-aliasing the resulting spotty image ("phosphene image"). We derive an equation concerning the correlations of pairs of phosphenes comprising the phosphene image, and show the relationship to the statistics governing the stochastic positioning. We present some examples where stochastic rendering assists the veridical perception of textures, and argue for its superiority as cf. ordered rendering. Our preliminary results suggest that it may be perceptually effective to manufacture disordered arrays of stimulating electrodes for intraocular implantation.

I. INTRODUCTION

IN the primate retina, cones are arranged in a slightly disordered mosaic [1]. Therefore, the light sourced and reflected by real-world objects is rendered on to the visual sense organ in a disordered fashion. Some workers hypothesize that the disordered nature of the retinal image is a central mechanism in the remarkable quality of human vision [2], specifically, its release from aliasing, which one would expect when a fine mosaic (like the cone mosaic) samples an even finer signal (for more discussion, including the role of eye optics, see [3]).

The retinal image and the phosphene image (see Fig. 1) bear similarity, albeit they operate at vastly different scales – seconds of visual arc versus degrees. This latter image is presently of interest to a number of vision researchers (e.g., [4], [5], [6]) as it serves as a visual model of electrical stimulation of the retina, a technique that may restore vision to those profoundly blinded through diseases affecting photoreceptors (for review see [7]). Studies involving the phosphene image have examined the recognition of (phosphenized) faces, reading speeds of (phosphenized) text, and the acuity that the phosphene image affords observers. This approach is analogous to the acoustic modeling of the cochlear implant, wherein mixed bands of colored noise are played to normally hearing listeners as a means of investigating speech processing strategies and implant candidacy criteria (e.g., see [8]). As shown in Fig. 1, the phosphene image is typically

comprised of relatively few phosphenes (usually on the order of 100), and is therefore subject to aliasing. Put differently, fine real-world detail is prone to manifest in the phosphene image as spurious low-frequency detail.

With the above in mind, the present authors have become interested in stochastic rendering and its use as an anti-aliasing mechanism, the aim being to improve perceptual outcomes concerning the phosphene image. In the next section, we derive an expression showing that correlations between phosphene pairs comprising the phosphene image tend to be cancelled by stochastic rendering. This, of course, depends on the statistics governing the rendering positions of phosphenes, and is generally more pronounced for high spatial-frequencies of the real-world scene being represented. In the Discussion, we take up the ramifications of all this for the perception of textures.

II. STATISTICAL ANALYSIS OF THE DISORDERED PHOSPHENE IMAGE

The covariance of any two phosphenes contained in the phosphene image is given by

$$\text{cov}(X, Y) = E(XY) - E(X)E(Y), \quad (1)$$

where $E(\cdot)$ denotes the expectation operator, and X and Y are random variables denoting the levels of activation of the two phosphenes. Then, consider a one-dimensional phosphene image, that is, a colinear array of phosphenes that, as opposed to assuming regular mosaic positions (that occur periodically with frequency $1/T_p$), are slightly (colinearly) displaced according to the probability density function (p.d.f.) denoted as $w(x)$. Then, the separation of any near neighboring phosphenes is governed by the sum of two independent random variables (see [10, p. 189]), that is, the p.d.f.

$$w'(x) = w(x) * w(x - T_p), \quad (2)$$

where the asterisk denotes convolution. If the one-dimensional phosphene image were driven by a cosinusoid with frequency f_s (it is convenient to think of the cosinusoid as an "underlying" texture for representation in the phosphene image), then by way of Eqs. (1) and (2) above, the covariance of near neighbors in the phosphene image is given by

$$\text{cov}(X, Y) = 1/8 \int_{-\infty}^{\infty} w'(x) \cos(2\pi f_s x) dx, \quad (3)$$

where $\cos(2\pi f_s x)/8$ is the mean-removed autocorrelation function of all arbitrarily shifted cosinusoids with frequency

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

D. S. Taubman is with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW, Australia, 2052.

G. J. Suaning is with the School of Engineering, University of Newcastle, Loughan, NSW, Australia, 2308.

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.

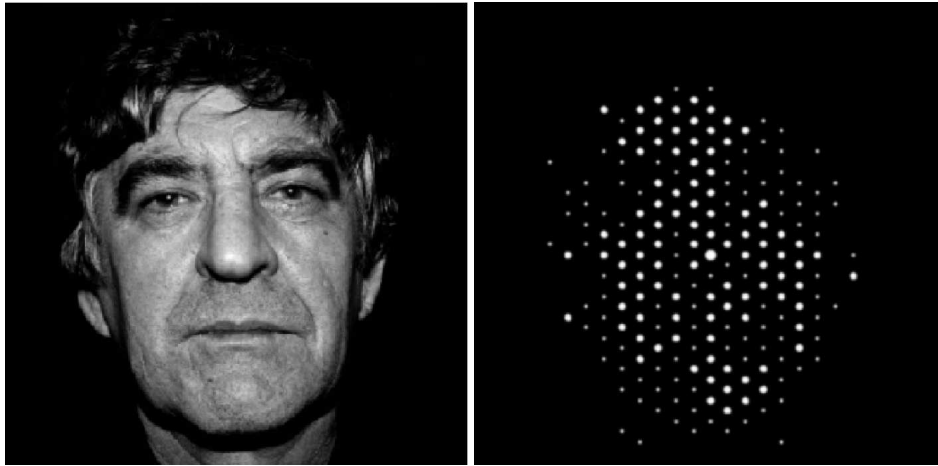


Fig. 1. An example phosphene image (right panel) – a visual model of electrical stimulation of the retina – representing a real-world scene (left panel). Microelectronic retinal prosthesis proposes to render luminous spots (so-called phosphenes), that may be modulated, in the visual field of the implant recipient. Together, phosphenes comprise the phosphene image. (Left panel image source: [9].)

f_s . Then, f_s in Eq. (3) may be varied, giving a function in frequency, f , that describes the covariance of near neighbors at various driving frequencies,

$$K(f) = \frac{1}{8} \int_0^\infty w'(x) \cos(2\pi f x) dx. \quad (4)$$

The reader will appreciate that $K(f)$ is one-sixteenth the cosine transform of $w'(x)$, that is,

$$\begin{aligned} K(f) &= \frac{1}{16} \mathcal{F}_c\{w'(x)\} \\ &= \frac{1}{16} \mathcal{F}\{w'(x) + w'(-x)\}, \end{aligned} \quad (5)$$

where $\mathcal{F}_c\{\cdot\}$ and $\mathcal{F}\{\cdot\}$ denote the cosine and Fourier transforms (FTs) respectively. The parameters for $K(f)$ are, firstly, k , the adjacency of the phosphenes in question ($k = 1$ for near neighbors, $k = 2$ for near neighbors but one, etc.), and, secondly, $w(x)$, the p.d.f. that governs the stochastic rendering. Since the variance of a sinusoid with amplitude A is given by $A^2/2$, then multiplication by a factor of eight for Eq. (5) yields the correlation coefficient (see [11, pp. 129–133]), which we will denote $\rho(f; k, w(x))$, on the range ± 1 .

By way of example, we now consider a specific class of phosphene image wherein phosphenes are rendered in positions according to the uniform p.d.f. $w(x) = \Pi(x/a)/a$, for a such that no two phosphenes overlap, where $\Pi(x)$ denotes a unit-area square pulse on support $[-1/2, 1/2]$. Then, from Eq. (2),

$$\begin{aligned} w'(x) &= w(x) * w(x - kT_p) \\ &= \frac{1}{a^2} \Pi(x/a) * \Pi(x/a - kT_p) \\ &= \frac{1}{a^2} \Lambda(x/a - kT_p), \end{aligned} \quad (6)$$

where $\Lambda(x)$ denotes the triangle function of unity height on support $[-1, 1]$. Then, the covariance function is the FT of $\Lambda(x/a - kT_p)/a^2 + \Lambda(kT_p - x/a)/a^2$.

This is depicted in Fig. 2. Note that as neighbor number, k , increases, that is, as replicas of $w'(x)$ are more divergent, the frequency of oscillation of $\rho(f; k, w(x))$ increases. As $a \rightarrow 0$, where in the limiting case phosphenes take their mosaic positions and each replica of $w'(x)$ is in fact a Dirac function, $\rho(f; k, w(x))$ oscillates indefinitely between ± 1 . This latter result indicates complete linear dependence arises between k -th neighbors when the phosphene image represents fine detail, that is, it indicates that the moiré effect-[12] is manifest in the image. We take up this point in the Discussion.

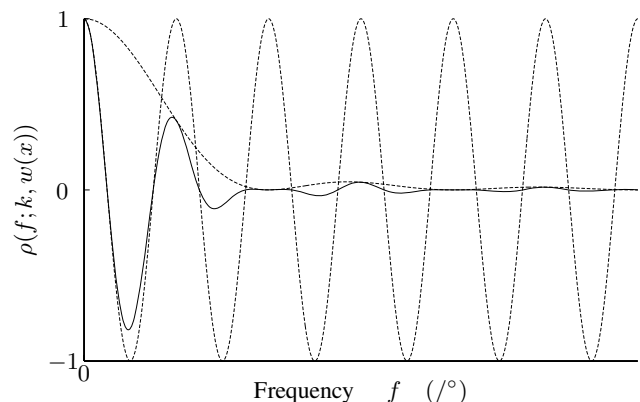


Fig. 2. The correlation, ρ , of k -th neighboring phosphenes (solid line) in the (uniformly) disordered phosphene image as a function of frequency, f , of the underlying stimulus. This is effectively the product of the two (dashed) envelopes shown. As neighbor number, k , increases, the frequency of the cosinusoidal envelope increases. The first zero of the $\text{sinc}^2(\cdot)$ envelope relates to the interval of the jitter, a ; as a increases, the zero decreases, and vice versa.

III. DISCUSSION

In the previous section, we derived an expression for the correlations between pairs of phosphenes as a function of f , the spatial-frequency component of the real-world scene being represented. We showed how this expression is affected

by the statistics that govern the jitter of phosphenes, and by k , the adjacency of phosphenes comprising the pair. This allows some predictions as to the appearance of the phosphene image. Firstly, the moiré effect tends to be annihilated by stochastic rendering. Secondly, textures involving fine detail take on the qualitative appearance of noise, as opposed to spurious structures.

Figure 3 provides two examples of the effect of stochastic rendering. In Fig. 3(a), a supra-Nyquist sinusoidal grating (top) is represented on an ordered phosphene mosaic (middle) and a disordered phosphene mosaic (bottom). Note that spurious structure is manifest (lower spatial-frequency and changed orientation) on the ordered mosaic (middle), whilst the entire disordered mosaic (bottom) takes on the appearance of noise. Despite the fact that a sine wave is an unlikely natural scene, it bears similarity to numerous features of the built environment which have important bearing on subject mobility, e.g., a flight of stairs. In Fig. 3(b), a Brodatz [13] texture (top) is similarly represented on both ordered (middle) and disordered (bottom) phosphene mosaics. Note how the ordered phosphene mosaic (middle) takes on the appearance of a composite surface (larger phosphenes appear to clump together). The disordered phosphene mosaic (bottom), on the other hand, takes on the uniform appearance of noise, which allows it to be segmented by the human visual system as a textural whole.

All of the foregoing begs the question, Why not bandlimit (low-pass filter) the real-world scene prior to sampling so as to avoid aliasing? Work in our lab shows that phosphene image observers typically scan the phosphene mosaic over the real-world scene and integrate phosphene outputs over time (using visual memory) so as to effect an increased sampling density (e.g., see the companion paper in this volume by S. C. Chen *et al.*). This being the case, to bandlimit the scene would be at odds with subjects' visuomotor behaviors (scanning), effectively decreasing the capacity of the (already impaired) information channel that is the phosphene image. Therefore, anti-aliasing measures that avoid low-pass filtering are of interest.

These results suggest that if a retinal prosthesis renders phosphenes in the visual field of the implantee in a disordered fashion then, all other things being equal, this is not perceptually disadvantageous. Rather, disorder affords perceptual benefits. It follows that, if an intraocular electrode array allows deterministic access to the visual field (that is, the layout of electrodes is faithfully reproduced by the layout of phosphenes in the visual field), then disordered arrays should be manufactured for implantation, as opposed to the ordered arrays (rectangular and hexagonal) that appear in the literature [14][15][16].

REFERENCES

- [1] J. Hirsch and W. Miller, "Does cone positional disorder limit resolution?" *Journal of the Optical Society of America A—Optics, Image Science and Vision*, vol. 4, pp. 1481–1492, 1987.
- [2] J. Yellott, "Spectral consequences of photoreceptor sampling in the rhesus retina," *Science*, vol. 221, pp. 383–385, 1983.

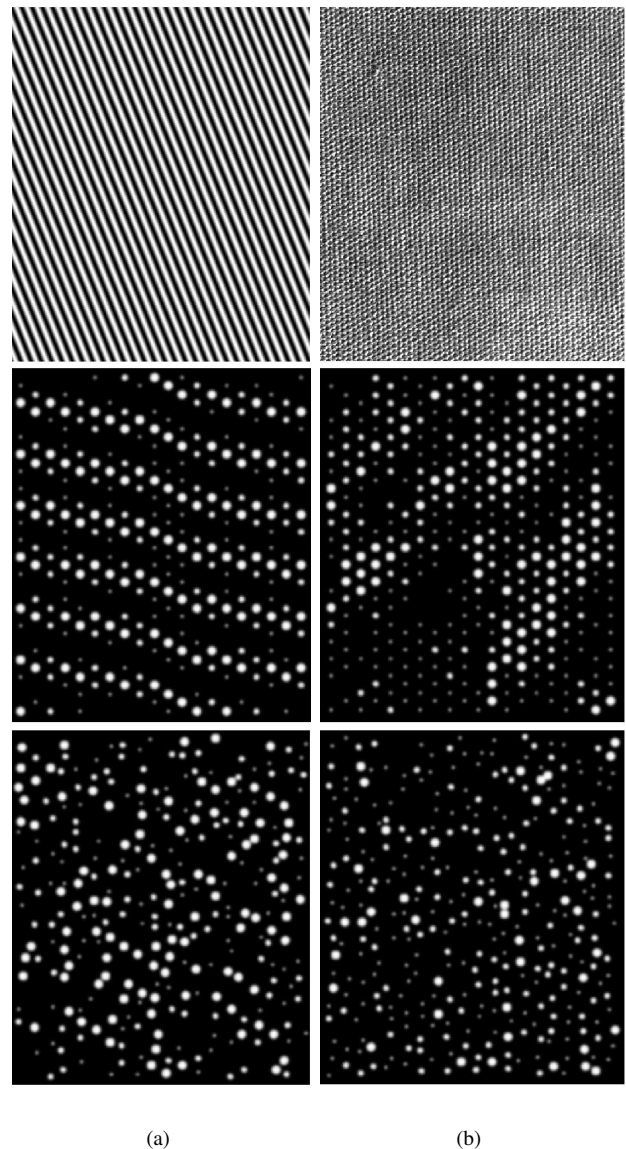


Fig. 3. Two examples of the effect of stochastic rendering, (a) a supra-Nyquist sinusoidal grating, and (b) a Brodatz [13] texture. In both cases, the top panel shows the original image, the middle panel shows the image rendered on an ordered phosphene mosaic and the bottom panel shows the image rendered on a uniformly disordered phosphene mosaic. Note that the ordered representations (middle panels) manifest spurious structure, whereas the disordered representations appear like noise (correlations between phosphene pairs are cancelled).

- [3] R. Navarro, P. Artal, and D. Williams, "Modulation transfer function of the human eye as a function of retinal eccentricity," *Journal of the Optical Society of America A—Optics, Image Science and Vision*, vol. 10, pp. 201–212, 1993.
- [4] 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.
- [5] 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.
- [6] 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.

- [7] E. Zrenner, "Will retinal implants restore vision?" *Science*, vol. 295, pp. 1022–1025, 2002.
- [8] R. Shannon, F.-G. Zeng, V. Kamath, J. Wygonski, and M. Ekelid, "Speech recognition with primarily temporal cues," *Science*, vol. 270, no. 5234, pp. 303–304, 1995.
- [9] P. Hancock, "Psychological Image Collection at Stirling," WWW: <http://pics.psych.stir.ac.uk>, accessed 1 July, 2003.
- [10] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, Tokyo, 1965.
- [11] H. Cramér, *The Elements of Probability Theory*. Wiley, 1955.
- [12] L. Spillmann, "The perception of movement and depth in moiré patterns," *Perception*, vol. 22, pp. 287–308, 1993.
- [13] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. Dover Publications, New York, 1966.
- [14] M. Humayun, J. Weiland, G. Fujii, R. Greenberg, R. Williamson, and J. L. et al., "Visual perception in a blind subject with a chronic microelectronic retinal prosthesis," *Vision Research*, vol. 43, pp. 2573–2581, 2003.
- [15] C. Merritt and B. Justus, "Fabrication of microelectrode arrays having high-aspect-ratio microwires," *Chemistry of Materials*, vol. 15, pp. 2520–2526, 2003.
- [16] M. Scheuttler, S. Stiebs, B. King, and G. Suaning, "Fabrication of implantable microelectrode arrays by laser-cutting of silicone rubber and platinum foil," *Journal of Neural Engineering*, vol. 2, pp. S121–S128, 2005.