A Cortical Integrate-and-Fire Neural Network Model for Blind Decoding of Visual Prosthetic Stimulation

Calvin D. Eiber, *Student Member, IEEE,* John W. Morley, Nigel H. Lovell, *Fellow, IEEE*, Gregg J. Suaning, *Senior Member, IEEE*

***Abstract*— We present a computational model of the optic pathway which has been adapted to simulate cortical responses to visual-prosthetic stimulation. This model reproduces the statistically observed distributions of spikes for cortical recordings of sham and maximum-intensity stimuli, while simultaneously generating cellular receptive fields consistent with those observed using traditional visual neuroscience methods. By inverting this model to generate candidate phosphenes which could generate the responses observed to novel stimulation strategies, we hope to aid the development of said strategies *in-vivo* before being deployed in clinical settings.**

I. Introduction

In the development of visual and other sensory neural prosthetics, the general architecture of the stimulus-response experiment is used to develop, test, and refine new kinds of electrode-tissue interfaces or stimulation strategies. Stimulus-response experiments which enable the researcher to discover how a neuron or population of neurons represents or encodes information are particularly useful for developing our general understanding of sensation. These experiments permit us to use the experimentally observed encoding, which is a hidden or unknown function, to formulate a decoding scheme which is a compact abstract model of the functional role of the observed circuit. Such a compact model can then be used to inform other models or layers of abstraction, and also as a guide to engineering for clinical outcomes.

A white-noise stimulus, in particular, is a useful input stimulus for sensory or other afferent neural circuits. This is due to the fact that, under certain assumptions about linearity, reverse spike-time correlation may be used to determine the encoding mechanism of a neural circuit [1]. This technique is particularly valuable because it protects the encoding model from biases which may arise from the *a-priori* assumptions which must be made by a researcher using most other kinds of stimulus-response experiment.

Because clinically implanted neuroprostheses must satisfy a variety of engineering requirements not shared by typical neurophysiology laboratory equipment, such as miniaturization for the sake of surgical feasibility [2], these experiments generally cannot draw upon the full breadth of neurophysiological stimulation techniques. Specifically, in such experiments, the stimulus must be of a form which could be produced by the prosthesis under investigation or development. For performing simple measurements, such as determining the minimum stimulation threshold which produces a clear neural response (an important parameter in the engineering of such devices [3]), this limitation is not overly significant. However, for experiments which begin to probe at more subtle characteristics of the stimulus-response relationship, it becomes necessary to have an understanding of how different characteristics of the neural code represent different characteristics of a natural stimulus which would be perceived as identical to the presented prosthetic stimulus.

It is typically not possible to present both natural and prosthetic stimulation in the same experimental environment – after all, if the patient population for a sensory neuroprosthesis could perceive that stimulus naturally, it would obviate the need to undergo surgery to receive a neuroprosthesis. Therefore, the neural encoding cannot be directly observed or decoded. It is in "blind decoding" tasks such as these where computational models of neural pathways become particularly useful. In the interests of developing a model which may be used to assist with the task of blind-decoding cortical responses to visual prosthetic stimulation, we present a leaky integrate-and-fire neural network model, based in large part on the prior work of McLaughin, *et. al.* [4].

II. Methods

A. Computational Model of the Visual Cortex

The model which this paper presents has been tuned to match cortical spiking statistics observed during the testing of a suprachoroidal visual prosthesis [5], and scales readily to rapidly simulate responses to both artificial visual inputs and natural images. This model incorporates convergent feed-forward excitation (or inhibition) from a layer which models the combined effects of the retinal neural circuitry and the circuitry of the lateral geniculate nucleus (LGN), as well as spatially isotropic lateral excitation and inhibition within the simulated cortical layer, layer 4Cα of the primary visual cortex. All modelling and analysis was conducted in Matlab (The Mathworks, Inc, Massachusetts, USA), using built-in packages wherever possible.

In the presented model, each cortical neuron is modeled as a single membrane voltage (normalized to be dimensionless) and associated membrane conductances at a point within the two-dimensional extent of the cortical layer. The vector of points which describes the cortical neuron centers is chosen randomly from within a 1×1 mm region of cortex and then relaxed by Lloyd's algorithm [6]. This results in cortical mosaics similar to the example shown in Fig. 1A. Roughly 25% of the cortical population was simulated as being inhibitory (noted as $s = I$, marked with squares in Fig. 1A),

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 (BV A).

C. D. Eiber, G. J. Suaning and N. H. Lovell are with the Graduate School of Biomedical Engineering, University of New South Wales, Sydney, Australia. (Phone: +61 2 9385 3911; fax: +61 2 9663 2108; e-mail: c.eiber@unsw.edu.au).

J. W. Morley is with the School of Medicine, University of Western Sydney, Campbelltown, Australia.

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

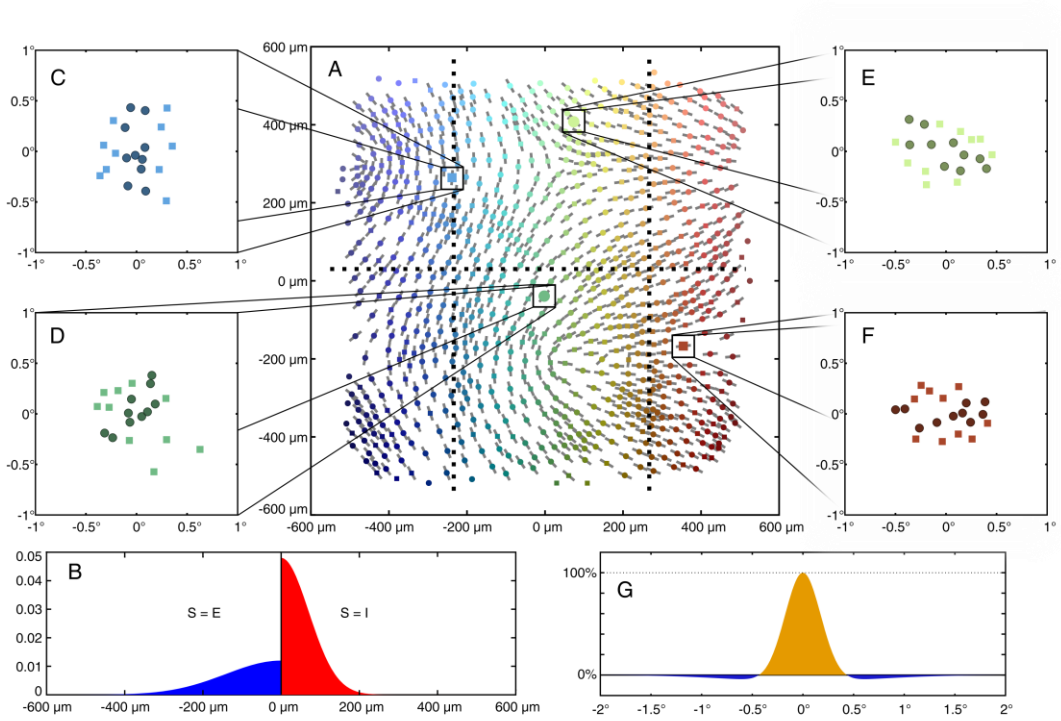

Figure 1. Example of cortical mosaic (A) with orientation preference, along with spatial cortical-cortical relationships (B), selected sample retinal mosaics (C-F), and spatial extent of retinal receptive fields (G). *(B and G are each normalized to unit integral).*

which provided a hyperpolarizing input to nearby cells [4]. The remainder were modeled as excitatory (noted as $s = E$, marked with circles in Fig. 1A), which provided a depolarizing input to nearby cells. Both excitatory and inhibitory cells may be depolarized by input from the LGN. The rate of change of the membrane voltage V of each cell is

$$\frac{dV}{dt} = -g_R(V - V_R) - g_I(V - V_I) - g_E(V - V_E), \quad (1)$$

in which g_R is the rest (or leakage) conductance (here, 50 s^{-1} [4]) and V_R is the resting potential; g_I is a variable inhibitory conductance and V_I is the corresponding inhibitory potential; and g_E is a variable excitatory conductance and V_E is the corresponding excitatory potential. When the membrane voltage exceeds the spike firing threshold, V_P, a spike on that neuron is registered and the neuron is reset to the resting potential V_R on the next time-step, which simulates the all-or-nothing nature of neural spiking. The membrane voltage itself is normalized such that $V_P = 1$ and $V_R = 0$ (in units of V/V), which gives $V_I = -2/3$ and $V_E = 14/3$ [4]. The conductances g_I and g_E are themselves given, for the j^{th} cortical neuron, by

$$g_{Ij} = g^I_{base} + S_{s,I} \sum_n^{s\in I} \left(\sigma^I_{jn} \sum_k^{K_n} \left(T_I(t - t_{n,k}) \right) \right) \quad (2)$$

$$g_{Ej} = g^E_{base} + g_{LGN} + S_{s,E} \sum_n^{s\in E} \left(\sigma^E_{jn} \sum_k^{K_n} \left(T_E(t - t_{n,k}) \right) \right)$$

in which $t_{n,k}$ is the time at which the k^{th} spike on the n^{th} neuron occurred; $T_s(\Delta t)$ is the time-course of inhibitory ($s = I$) or excitatory ($s = E$) cortical coupling, given by (4) (found in the appendix) and shown in Fig. 2A; K_n is the number of spikes which have previously occurred on the n^{th} neuron; σ^s_{jn} is the connectivity between the n^{th} neuron and the j^{th} neuron, given by (5) and shown in Fig. 1B for both $s = I$ and $s = E$ (calculated based solely on the distance between the pair on the cortex), and $S_{a,b}$ is the cortical coupling strength, where a is the polarity of the j^{th} neuron and b is the polarity of the n^{th} neuron. The input to this neuron from the neurons in the LGN is modelled by g_{LGN}, given below in (3). Stochastic input from other regions of the cortex, which determines the baseline rate of spike generation, are modelled as uniform probability distributions of amplitude $g^{I/E}_{base}$, which is treated as an adjustable parameter of the model.

Each cortical neuron receives convergent input from a number of neurons in the LGN. For the j^{th} neuron, this input is modelled by the equation

$$g_{LGN} = \sum_n^{N_j} \left(T_{LGN}(\alpha t) * I_0 \iint I(\vec{X}, t) f_s(\|\vec{X} - \overrightarrow{X_n}\|) d\vec{X} \right), \quad (3)$$

in which N_j is the number of receptive fields which project to the j^{th} neuron (fixed at 20 [4, 7], with an equal number of

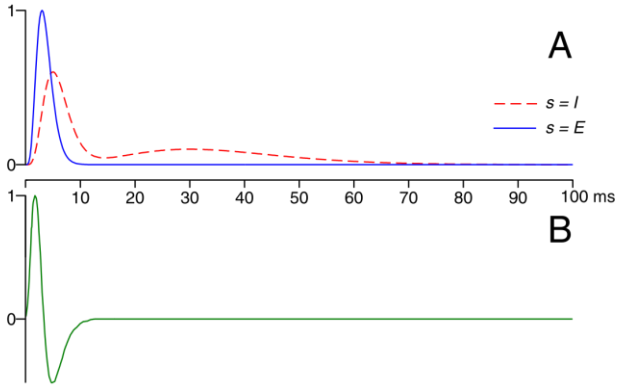

Figure 2. Time Dynamics for cortical coupling (A) and LGN input (B).

on-centre and off-centre receptive fields); T_{LGN} is the time-course of the LGN response to an impulse (which may be scaled in time by a factor of α), shown in Fig. 2B, which is convolved with the overall instantaneous response of the retinal receptive field; I_0 normalizes the input intensity to a unit input, $I(\vec{X},t)$ is a function which gives the intensity of the input at a location $\vec{X}$ on the retina at time t; and $f_s(x)$ gives the spatial sensitivity of an on-centre or off-centre receptive field, given in (6) and shown in Fig. 1B.

Equation (3) is a simplified model of the precortical components of the visual pathway which incorporates both the temporal filtering applied by the cells of the LGN and the centre-surround organization which is implemented lower in the visual pathway for natural input, although more detailed models of the precortical visual pathways may be substituted in future work. The receptive field centres which project to each neuron in the LGN are generated by a method similar to the procedure which generates the cortical mosaic; several examples are shown in fig. 1B. The orientation selectivity for each neuron is inherited from the pinwheel organization of the hypercolumns of the cortex [8], outlined in Fig. 1A by the dotted lines superimposed on the cortical mosaic.

B. *Experimental In-Vivo Procedures*

All surgical and experimental procedures were reviewed and conducted with approval from the UNSW Animal Care and Ethics Committee.

Our *in-vivo* experimental methods have been published previously [5]. To summarize, anesthesia was induced and maintained in normally sighted cats for the duration of the multi-day experiment. A custom stimulating array consisting of 24 platinum disk electrodes [9] was introduced into the suprachoroidal space, positioned 1 mm lateral to the optic disk, and verified by fundus imagery.

Access to the visual cortex was achieved by performing a craniotomy and durotomy. The primary and secondary visual cortices were mapped with a platinum ball electrode while stimulating the retina with suprathreshold stimuli. Once the cortical area of maximum response was located, a 100 channel (10x10) penetrating electrode array (Blackrock Microsystems, Utah, USA) was inserted, and data from all electrodes was recorded simultaneously (Tucker-Davis Technologies, Florida, USA).

The stimulating electrodes were then stimulated with respect to a monopolar platinum return electrode placed beneath the eyelid. Current-controlled biphasic stimuli (500 µs/phase) at varying amplitudes were presented in random order, and the cortical response recorded; however, for the purposes of fitting our model to our recorded data, only sham stimuli (zero amplitude) and stimuli delivered at the greatest amplitude setting (500 µA) were used. During offline processing, stimulus artefacts were removed from the recording, and the signal was band pass filtered from 300 Hz to 5000 Hz. Pre-stimulus times were used to calculate the baseline RMS noise content of the signal. Spikes were identified as the times at which the recording exceeded 3.6 times the RMS noise level [5], which lead to low baseline spike-rates and large stimulation spike-rates. For each recording from one of the 10 chosen stimulation sites, the "best cortical electrode" (BCE) was identified in post-processing as the recording electrode site at which the half-maximal trial-averaged spike-rate could be achieved with the minimum stimulus amplitude. For some stimulation sites, the BCE was not stable over multiple sessions; for the purpose of model fitting, only data recorded at the BCE for a given session was used for model training.

III. Results

Our model is capable of consistently matching experimentally observed spiking statistics to a high degree of fidelity, as shown by the overlay of experimental data and simulated data displayed in Fig. 3. We found that the width of the LGN kernel T_{LGN} for the response to electrical stimuli was much less than previously reported values [4], and the other adjustable parameters of the model took the following values:

TABLE I. Parameters of Simulation

Symbol	Description	Value (units)
$S_{I,I/E}$	Cortical synaptic strength, I→I & , I→E	7.6 (s^{-1})
$S_{E,I}$	Cortical synaptic strength, E→I	1.5 (s^{-1})
$S_{E,E}$	Cortical synaptic strength, E→E	0.8 (s^{-1})
$g_{base}^{I/E}$	Stochastic baseline activity	0 – 88 (s^{-1})
I_0	Scaling factor for input	220 (s^{-1})
α	Scale of LGN time dynamic	15×

We fit the observed maximum monopolar response to an input consisting of a unit-intensity uniform flash of duration 1 ms presented to the model once it had reached steady-state. Using the values in table 1, the maximum values of the cortical inhibitory conductances g_I reached approximately 300 s^{-1} at around 10 ms post-stimulation, and the maximum values of the excitatory conductances g_E (due to both cortico-cortical connections and input from the LGN) reached approximately 4000 s^{-1}, for brief durations. When individual neurons were isolated from other neurons in the cortex by setting each $S_{I/E,I/E}$ to 0, the maximum excitatory conductance g_E due solely to input from the LGN was on the close order of the excitatory conductances reached under the full model, indicating that the cortical component of the excitatory conductance is less than the purely cortical inhibitory

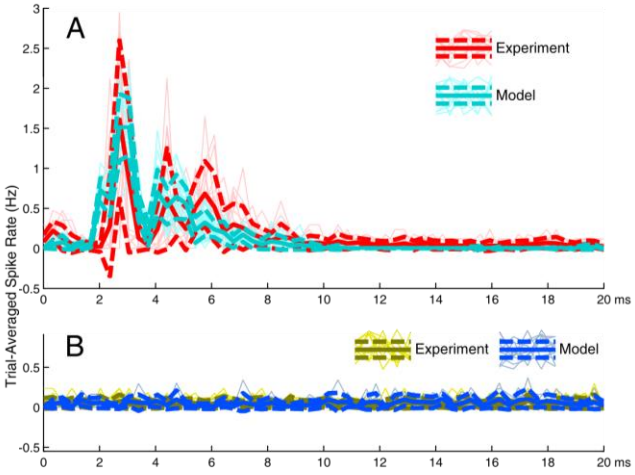

Figure 3. Comparison of recorded and simulated stimulus response (A) and sham response or baseline (B). RMSSD = 1.9. *(Raw histograms shown by faded lines; mean ± S.D. shown by thick lines).*

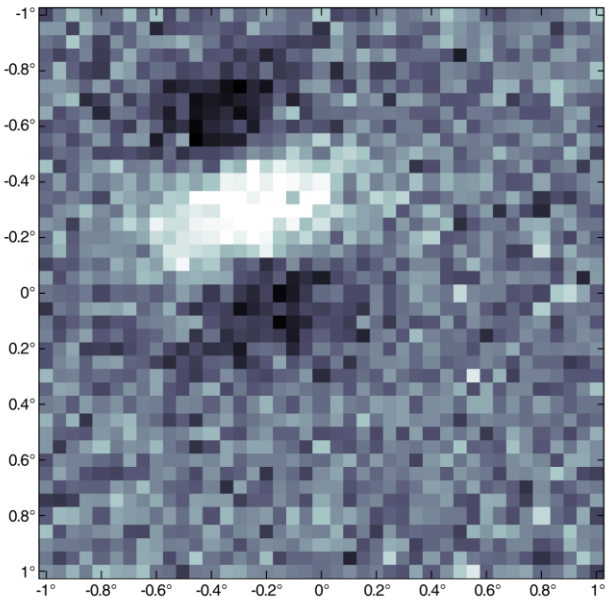

Figure 4. Typical example recovered visual receptive field at t = 1 ms pre-spike over 30 s of simulation. Greyscale shows spike-triggered average intensity of each pixel in the visual field, scaled such that white and black represent ±3 std. dev. of the time-averaged intensities from the mean.

conductance, which would be expected. The model fit was assessed using the RMS standard deviation [11] averaged over the timecourse of the simulation. On average, the model fits the stimulus data to within 1.9 standard deviations.

When the model was presented with uniformly distributed white-noise stimuli (refreshed at 200 Hz), the receptive fields of the cortical neurons could be reconstructed using reverse spike-time correlation, as shown by the example reconstructed field shown in Fig. 4. These receptive fields are consistent with the distribution of physiological receptive fields found in simple cells of the visual cortex [8].

IV. Conclusion

The model which we have presented here is both computationally efficient and able to consistently reconstruct the observed distributions of cortical activity in response to prosthetic stimulation. This model suggests explanations for nuances in the distributions of cortical activity which were previously unexplained, which we now believe to arise from the cortical responses to both activation and suppression of retinal ganglion cells by the electrical stimulus. By exploring this model and its permutations, we hope in the future to develop a more detailed picture of precisely how varying stimulation strategies interact with retinal circuits in order to develop high-fidelity techniques which may be clinically applicable to the restoration of visual sensation to patients with retinal degenerative diseases.

Appendix

The time-course of inhibitory ($s = I$) or excitatory ($s = E$) cortical coupling (in Hz) is given by

$$T_s(t) = \frac{t^5}{\tau_s^6} * e^{-t/\tau_s}, \tag{4}$$

in which the time-course of activation is controlled by the time constant τ_s, which takes the value 3/5 ms for ($s = E$) and the values 1 ms and 6 ms for ($s = I$). Peaks occur at $5\tau_s$. Equation (4) is shown graphically in Fig. 2A.

The spatial extent of inhibitory ($s = I$) or excitatory ($s = E$) cortical coupling for two neurons separated by a distance x is given by the Gaussian

$$\sigma_s(x) = \frac{(\overline{\Delta x})^2}{\pi L_s^2} e^{-x^2/L_s^2}, \tag{5}$$

in which the median distance between cortical neurons is $\overline{\Delta x}$ (a constant model parameter) and the extent of the arbours of the different cortical cells is represented by the length constant L_s, which takes the value 200 µm for ($s = E$) and the values 100 µm for ($s = I$) [4]. Equation (5) is shown graphically in Fig. 1B.

The sensitivity of a retinal receptive field at a distance x from the receptive field center is given by the difference of Gaussians

$$f_s(x) = \frac{17}{r_c^2} e^{-\frac{x^2}{r_c^2}} - \frac{16}{r_s^2} e^{-\frac{x^2}{r_s^2}}, \tag{6}$$

in which the extent of the center of a receptive field r_c is 0.25° and the extent of the surround of a receptive field r_s is 1° [12]. Equation (6) is shown graphically in Fig. 1G.

References

[1] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational And Mathematical Modeling of Neural Systems*: Massachusetts Institute of Technology Press, 2005.

[2] T. Stieglitz, "Manufacturing, assembling and packaging of miniaturized neural implants," *Microsystem Technologies,* vol. 16, pp. 723-734, 2010/05/01 2010.

[3] C. D. Eiber, N. H. Lovell, and G. J. Suaning, "Attaining higher resolution visual prosthetics: a review of the factors and limitations," *J Neural Eng,* vol. 10, p. 011002, 2013.

[4] D. McLaughlin, R. Shapley, M. Shelley, and D. J. Wielaard, "A neuronal network model of macaque primary visual cortex (V1): Orientation selectivity and dynamics in the input layer 4Cα," *Proceedings of the National Academy of Sciences,* vol. 97, pp. 8087-8092, July 5, 2000 2000.

[5] P. B. Matteucci, S. C. Chen, C. Dodds, S. DokosNigel, H. Lovell, and G. J. Suaning, "Threshold analysis of a quasimonopolar stimulation paradigm in visual prosthesis," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 2997-3000.

[6] S. Lloyd, "Least squares quantization in PCM," *Information Theory, IEEE Transactions on,* vol. 28, pp. 129-137, 1982.

[7] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J Physiol,* vol. 160, pp. 106-54, Jan 1962.

[8] D. H. Hubel, *Eye, brain, and vision*. New York: Scientific American Library : Distributed by W.H. Freeman, 1988.

[9] C. W. D. Dodds, M. Schuettler, T. Guenther, N. H. Lovell, and G. J. Suaning, "Advancements in electrode design and laser techniques for fabricating micro-electrode arrays as part of a retinal prosthesis," in *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 636-639.

[10] R. J. Tusa, L. A. Palmer, and A. C. Rosenquist, "The retinotopic organization of area 17 (striate cortex) in the cat," *J Comp Neurol,* vol. 177, pp. 213-235, 1978.

[11] C. D. Schunn and D. Wallach, "Evaluating goodness-of-fit in comparison of models to data," *Psychologie der Kognition: Reden and Vorträge anlässlich der Emeritierung von Werner Tack,* pp. 115-154, 2005.

[12] T. W. Troyer, A. E. Krukowski, N. J. Priebe, and K. D. Miller, "Contrast-invariant orientation tuning in cat visual cortex: thalamocortical input tuning and correlation-based intracortical connectivity," *The Journal of neuroscience,* vol. 18, pp. 5908-5927, 1998.