Physiologically Based Analysis of Cochlear Implant Representations

J.B. Laflen\textsuperscript{1,2}, T.M. Talavage\textsuperscript{1,2}, P.M. Thirukkonda\textsuperscript{2}, M.A. Svirsky\textsuperscript{3}

\textsuperscript{1}Department of Biomedical Engineering, Purdue University, West Lafayette, IN
\textsuperscript{2}School of Electrical & Computer Engineering, Purdue University, West Lafayette, IN
\textsuperscript{3}Department of Otolaryngology-HNS, Indiana University School of Medicine, Indianapolis, IN

Abstract—A method is presented for analyzing cochlear implant stimulations and typical representations used in simulations. Filtered “white-noise” bands are modulated using sinusoids, representing differing stimulation channels. These representations, along with their corresponding envelopes, are used to generate neural activation patterns (NAPs), which represent “normal-hearing” responses in the auditory nerve to these stimuli. Additionally, NAPs are generated to represent the neural activity induced by cochlear implant stimulation strategies, assuming exponential rolloff from the electrodes. The mean squared error is measured between NAPs both directly, and after compensation for perceptual resolution. Results suggest that the noise-band approximation of the CIS implant signal actually has more in common with the original source than with the implant stimulation patterns.

Keywords—electrical stimulation, implants, simulation

I. INTRODUCTION

Over 28 million adults are deaf or hearing impaired, and as many as three in every thousand children are born with a congenital hearing loss or will acquire one at an early age. Approximately 10\% of patients with sensori-neural hearing loss are eligible for cochlear implants [1]. Modern cochlear implants consist of an external detection microphone, a speech processor for determining stimulation parameters, and an n-channel electrode array for delivering the stimulus. All processing strategies lead to stimulation by the electrode array, which in turn leads to activation in the auditory nerve.

There is a wide variation in success across implant users: some patients achieve near-normal speech perception levels while others obtain only some lipreading improvement. The desire to deliver more information to the auditory nerves has motivated innovations in electrode stimulation strategies. One such strategy is Continuous Interleave Sampling (CIS), which divides the instantaneous power spectrum of the external signal into n frequency bands (corresponding to the n electrode array channels) and uses the relative energy in each band to modulate activation of each implant channel.

Analysis of performance variations across implant users requires comparisons amongst stimulation strategies and the “normal-hearing” case. Such comparisons can be made through psychophysical experiments by presenting acoustic signals that simulate implant activation to “normal-hearing” subjects. These simulations are commonly generated using noise-bands. The effects of implants and noise-bands upon perception are not easily known, largely due to complexities within the auditory system [2], [3], [4]. This paper presents one method for determining these perceived effects while accounting for the complexities of human hearing.

II. METHODOLOGY

Let $y(n)$ be any input sound signal sampled with frequency $f_s$. If N tones at center frequencies $\{f_i(n)\}$ and with amplitudes $\{a_i(n)\}$ and phase $\{\theta_i(n)\}$ were selected to match the instantaneous spectrum of $y(n)$ (either by feature extraction or band-selection, e.g., CIS), then a sine-wave reconstruction of $y(n)$ is

$$x(n) = \sum_{i=1}^{N} a_i(n) \sin(2\pi f_i(n) \cdot n/f_s + \theta_i(n))$$

Let $n_{L,\xi}(n)$ represent white-noise, filtered with a low-pass ($f_c = \xi/2$) FIR filter and Hamming window. Then,

$$n_{B,\xi}(n) = n_{L,\xi}(n) \cdot x(n)$$

is a collection of noise-bands (bandwidth $\xi$) centered at $\{f_i(n)\}$, which commonly represents electrode activation at cochlear locations tonotopically associated with $\{f_i(n)\}$.

A neural activation pattern (NAP) represents activation in both place and time within the auditory nerve. We generate a NAP using the Auditory Image Model (AIM, [5]). In AIM, the NAP is actually a set of response transformations centered at various simulation frequencies that correspond to locations along the cochlea. We generate WAV files of $x(n)$, $y(n)$, and $n_{B,\xi}(n)$, and use AIM to generate corresponding NAPs: $X(f, n)$, $Y(f, n)$, and $N(f, n)$ respectively.

We generate NAP representations of CIS stimulation using the parameters $\{f_i(n)\}$, $\{\theta_i(n)\}$, $\{a_i(n)\}$, and $\xi$ from above and assuming a stimulation rate $f_{stim}$. Here, an exponentially decaying excitation pattern models the electrode stimulation, resulting in the NAP:

$$I_i(f, n) = a_i(n)e^{-2|f-f_i(n)|/\xi} \cdot \sin(2\pi f_{stim} \cdot n/f_s + \theta_i(n))$$

We sum across all N NAPs to generate a NAP representative of the cochlear activation due to CIS:

$$I(f, n) = \sum_{i=1}^{N} I_i(f, n)$$
Y, X, N_B,ξ, and I represent neural activation in the auditory nerve from their respective stimulus sources. We approximate the “perceptual difference” between these NAPs by applying a “perceptual transformation” (T_ξ[·]) to each NAP, and then computing the mean squared error:

\[
\Delta P_ξ(A,B) = \frac{1}{FN} \sum_{\{f\}} \sum_{\{N\}} |T_ξ [A] (f,n) - T_ξ [B] (f,n)|^2
\]

where \(\{F\}\) (size \(F\)) is the set of frequencies associated with location along the cochlea in the NAP. The “perceptual transformation,” \(T_ξ[·]\), is any transformation that weights the NAP to account for difference limens in intensity, spatial discrimination, and/or temporal rate.

Generally, \(ξ\) varies with center frequency to account for the logarithmic mapping of frequency to place in the cochlea. Thus, \(ξ\) is set to match, in frequency bandwidth, an approximate electrode resolution.

III. Results

Our results examine a simplified scenario involving the perceptual response to a single frequency excitation: \(N = 1, \{f_i(n)\} = \{2000 \text{ Hz}\}, \{a_i(n)\} = \{1\}, \{θ_i(n)\} = \{0\}, \) and \(ξ = 500 \text{ Hz}\) (which approximates 1 mm of stimulation resolution at 2 kHz). Using (1) and (3) we generated WAV files of \(x(n)\) and \(n_{B,500}(n)\) with \(f_s = 11025\) Hz and simulation time of 0 to 4 s, and generated NAPs using AIM with 100 frequency channels (in Equivalent Rectangular Bandwidths) between 100 Hz and 5 kHz (note that \(N = 1 \rightarrow y(n) = x(n)\), so a separate representation of \(y(n)\) and its NAP was unnecessary). Similarly, we generated a NAP representation of the CIS equivalent of \(y(n)\) using (5) with \(f_{stim} = 500\).

Fig. 1 shows decimated illustrations of the NAPs: [A] \(X(f,n)\); [B] \(N_{B,500}(f,n)\); and [C] \(I(f,n)\).

Using (6) we calculated \(\Delta P\) using two “perceptual transformations”: \(T_1[A(f,n)] = A(f,n)\), which asserts that all activation is equally weighted for perception, and \(T_2[A(f,n)] = \log_{10} A(f,n)\), which is motivated by

\[\begin{array}{ccc}
T_1[A(f,n)] = A(f,n) \\
\Delta P_1(X, N_{B,500}) & \Delta P_1(X, I) & \Delta P_1(N_{B,500}, I) \\
0.0043 & 0.0246 & 0.0246 \\
T_2[A(f,n)] = \log_{10} A(f,n) \\
\Delta P_2(X, N_{B,500}) & \Delta P_2(X, I) & \Delta P_2(N_{B,500}, I) \\
0.0074 & 0.0315 & 0.0316
\end{array}\]

IV. Discussion

The data in Table I show a smaller \(\Delta P\) between \(x\) and \(n_{B,500}\) than between the CIS stimulation and either \(x\) or \(n_{B,500}\), for both \(T_1\) and \(T_2\). This suggests that a noise-band is perceptually closer to a pure tone than the electrode stimulation that it commonly represents with “normal-hearing” subjects. Further, the data also suggest that the CIS stimulation does not closely mimic normal auditory activation, even in the case of a single frequency excitation centered on the stimulation location (i.e. with no basal shift).

V. Conclusion

In this paper, we have used the NAP to evaluate CIS stimulation as it relates to normal peripheral processing. Our results suggest that noise-bands may not be the “best” representation of implant stimulation, and also that there is a large perceptual difference between the original acoustic signal and its CIS implementation. This last point coincides with, and may partially explain, the extended post-implant time period that is usually required for cochlear implant patients to achieve asymptotic speech perception scores.

REFERENCES


\[\text{1} \text{All NAPs were normalized to values between 0 and 1 before computing } \Delta P.\]