STIMULUS SIGNAL ESTIMATION FROM AUDITORY-NEURAL TRANSDUCTION INVERSE PROCESSING

Habibzadeh V. Houshang*, Kitazawa Shigeyoshi

Graduate School of Science and Engineering, Shizuoka University 3-5-1 Johoku, Hamamatsu 432, JAPAN E-mail: houshang@cs.inf.shizuoka.ac.jp

ABSTRACT

Auditory models reverse processing techniques would have very useful applications in speech perception and auditory models evaluation. This paper examines how we can be benefit an Inner Hair Cell (IHC) model as a compression and envelope detection section, in the cochlear model inverse processing. Our proposed inversion method, combines the reverse of the Meddis's auditory neural transduction model with Lyon's cochlear model to estimate the input signal to the inner ear from its auditory nerve firings, with the acceptable quality. Since this method uses neural firings or cleft contents as an input and re-generates the original acoustic stimulus, it is useful with any system generating auditory neural firings. For example, using this method, we are able to estimate the stimulus signal of the Nucleus Cochlear Implant systems to investigate the transferred speech quality without using the real patients.

1. INTRODUCTION

Computational models of auditory peripheral processing have been popular for many years. But there have been less work on auditory models inversion, which could be used in auditory models evaluation and speech enhancement applications. In the first cochlear model inversion study, Daniel Naar [1] and Malcolm Slaney [2] reverse processed the Lyon's cochlear model, to resynthesis the original signal from it's Correlogram. Their primary target was to separate a sound from noisy background. Some earlier works outlines methods for evaluating auditory models; one common technique is to use the model as a pre-processor, or "front-end " to an automatic speech recognizer. The model is then considered to be improved when the recognition rate improves. However the result of such testing will inevitably be influenced by the recognizer characteristics themselves. An alternative possibility is to calibrate the model against human perceptual data, but problem of language bias and task differences make this difficult. Some earlier studies introduced the idea of *testing by* Fig.



1 Hybrid cochlear Model(reverse and forward).

resynthesis [3,4]. According to this idea, two different acoustic signals having the same auditory representation should sound equivalent in some important respect. It is believed that the extent of the perceptual differences between two such signal is a good measure of the model's quality. Thus, evaluation could be performed by listening tests without any necessity to consider the many extraneous factors due to an automatic recognizer. Following the works of Ray Meddis[5][11][12] in Mechanical to Neural transduction in the auditory nerve fibers, Malcolm Slaney and Daniel Naar [1][2] in Sound re-synthesis from a correlogram, we propose a method that combines an auditory neural transaction model with a suitable filter bank(i.e. cascade filter bank of Lyon's cochlear Model) to produce more realistic presentation of the human inner ear characteristics in forward and reverse models.

During the past three decades, a number of models of primary auditory fiber activity have been developed which generate a sequence of firings in time, in response to a stimulating waveform. The most recent one, Meddis model[5], is a simple and computationally efficient model for neuro-transduction in inner hair cell. So, we used this model as an automatic gain control (AGC) and half-wave rectification (HWR) section of Lyon's Cochlear model [6] for forward and reverse processing. The new hybrid model has two sections, Meddis inner hair cell and Lyon's filter bank (Fig. 1).

2. COCHLEAR MODELS

The cochlear band-pass filters, the half-wave rectification, and the automatic gain control are combined in the processing performed by the ordinary *cochlear model*.

2.1. Ordinary Cochlear Models Structure.

However, all cochlear models differ in their underlying assumptions and structures, but they share three primary characteristics [7].

- FILTERING : A broadly tuned cascade of filters to model the propagation of energy as waves on the basilar membrane (BM).
- DETECTION: which simulates the HWR characteristics of the IHC.
- COMPRESSION : which maps the widely varying sound input levels into a limited dynamic range of BM motion, IHC and auditory nerve (AN) fibers.

2.2. Meddis Hair Cell Model

In order to get better understanding of the IHC reverse processing section, here is the block diagram of the Meddis IHC forward model from his 1990 JASA paper [5]. The model can be viewed in terms of the production, movement, and dissipation of the transmitter substances in the region of the hair cell and the auditory nerve fibers The k(t) function is intend to reflect the permeability of the membrane. It is given as the direct function of the stimulus intensity S(t). The probability of spike can be found by multiplying cleft contents c(t) to a constant *h* at a given instance.

Prob(event) = hc(t)dt

3. COCHLEAR MODEL INVERSION

In order to estimate the original time sequence of the signal that entered the cochlear model from the auditory nerve firings, all sections in the cochlear model must be inverted.

3.1. Ordinary Cochlear Model Inversion

In the earlier inversion study[1][2], which was based on Lyons cochlear model , separate AGC, HWR, and filtering sections have been inverted.

3.2. IHC Inversion Process

In the hybrid model inversion (Fig. 1.), to estimate the original signal, Filtering and Meddis hair cell sections have to be inverted. The following section explains the necessary operations to estimate the input acoustic stimulus from IHC section(Fig.2.) only for a single channel.



 $\begin{array}{l} dq/dt = y(M-q(t)) + xw(t) \cdot k(t)q(t) \\ dc/dt = k(t)q(t) \cdot lc(t) \cdot rc(t) \\ dw/dt = rc(t) \cdot xw(t) \\ k(t) = gdt(S(t)+A))/(S(t)+A+B) \ for \ S(t)+A > 0 \\ k(t) = 0 \ for \ S(t)+A < 0 \end{array}$

Fig. 2. Meddis hair cell model, and it's differential equations.

Using Meddis model and his last parameter-set, we wrote programs with MatlabTM to calculate the cleft contents (or firing rates) and its reverse relations. Having c(t) (given or calculated from forward processing), we used equation (1) to calculate permeability function k(t). Substituting c(t_i) and calculated k(t_i) in equations (2) and (3), we compute the amount of the reservoirs $q(t_{i+1})$ and $w(t_{i+1})$ for the next time slice $t_{i+1}(t_{i+1}=t_i+1/SF)$. Then using equation (4), we compute the s(t_i) and start again with c(t_{i+1}). At the end of process, the regenerated signal has been stored in the vector S(t).

4. EXPERIMENTS

Here, we outline the result of reverse process for three experiments with different input stimulus. All experiments and calculations have been done with Matlabtm, Mathematicatm and Turbo Pascal programming under Windows 95tm, and SUN-OStm, using SUN-Spark Work stations and a Pentiumtm PC. To calculate the S(t) from given c(t), we used the following equations in our program with time step dt of 1/(Sampling Frequency).

$$k(t) = (c(t+1)+c(t)(rdt+ldt-1))/q(t)$$
(1)

$$q(t+1) = q(t) - k(t)q(t) + (xw(t) + y(M - q(t)))dt$$
(2)

$$w(t+1) = w(t) + (rc(t) - xw(t))dt$$
 (3)

$$S(t) = \frac{gAdt - k(t)(A+B)}{k(t) - gdt}$$
(4)

The following Meddis parameters have been used in the

forward and reverse processing:

A=5, B=300, y=5.05, g=2000, l=2500, r=6580, dt=1/SF, x=66.31, M=1.

The initial condition of system is determined from the equilibrium state of the system with no input: q=c(l+r)/k(t); w=cr/x where k(t)=gA/(A+B).

4.1. Pure Tone Sinusoidal Stimulus

In the first experiment, we used a 80 dB 1 kHz sinusoidal tone, as a stimulus signal S(t) for the Meddis model. The generated data in this step c(t) have been saved and used in the inversion process. At the IHC inversion process, we used c(t) as input, and we generate permeability function k(t), all reservoir contents, and stimulus signal S(t) (Fig. 3).



Fig. 3 Re-generated sinusoidal signal.

4.2. Sound (Music) Stimulus

In the second experiment, we extracted data from a 22.254 kHz, 8 bit, PCM WAV file. The data has been used in forward processing to generate neural firing rates. Here also we used the output c(t) to estimate the original signal, S(t). Figures 4, 5 and 6 show the original, compressed and regenerated signal using inversion process.

4.3. Speech Stimulus

In the third experiment, data was extracted from an 11 kHz speech WAV file. After adjustment of the zero line, input signal has been regenerated. Surprisingly, the estimated signal in the listening test, almost was not distinguishable from the original one. Figures 7,8 and 9 show the original and regenerated speech signal wave forms.

5. CONCLUSIONS

A cochlear model inversion method based on Meddis hair cell model has been defined. In this study, a simple and efficient method has been provided to achieve acceptable quality without using complex multiplicative and multistage AGC and HWR of the earlier work. Regarding to the outputs of the reverse processing, generated signal from various stimulus types are barely degraded. Regenerated positive part is almost the same as original one. We find this method less accurate in very high stimulus amplitude (Fig. 3, with distortion) and low sampling frequency (Fig. 8, with poor compression). In overall, the method shows its ability to re-generate any input stimulus to the inner ear with acceptable quality.



Fig. 4 Original stimulus PCM, 8bit WAV signal.(Ding! sound) over discrete time step dt(1/SF).



Fig. 5 Half wave rectified and compressed output of the forward processing, which defines the contents of the *synaptic cleft* area or c(t) in discrete time dt .



Fig. 6 HWR version of the estimated stimulus signal using IHC inversion process.(Time step=1/SF).



Fig. 7 Original stimulus WAV signal (word *Pan*), over discrete time step dt.



Fig. 8 Half Wave Rectified and compressed output of the forward processing, which defines the contents of the synaptic cleft area or C(t), over discrete time step dt (dt=1/SF).



Fig. 9 HWR version of the estimated speech signal.

REFERENCES

[1] Daniel Naar " *Sound resythesis from a correlogram*" Apple technical Rep. No.3 May 1993.

[2] Slaney et al. "Sound Analysis and Resynthesis using correlograms" Apple computer Inc. Dec. 1995.
[3] Damper, R.I., Bladon, R.A.W., Hukin, R.W. and Irvin, G.N.A (1987) "Resynthesis and matching experiments on auditory theory male/female normalization", Proc. XI Int. cong. Phonetic Sci., Tallinn, ESSR, 78-81.

[4] Hukin, R.W. (1988) "*Testing auditory models by resynthesis*", Unpublished thesis, University of Southampton.

[5] Rey Meddis et. al., "*Implementation details of a computational model of the IHC/AN synapse* " J. Acoust. Soc. Am. 87(4), April 1990.

[6] Lyon, R.F. (1982) "A Computational Model of Filtering, Detection, and Compression in the Cochlea". IEEE Transaction on acoustic, Speech, and Signal Processing, 12, 1282-1285.

[7] Malcolm Slaney, R. F. Lyon "On the importance of time- a temporal representation of sound" Visual

representation of speech signals, Wiley&Sons 1993.

[8] L.R. Rabiner and R.W. Schafer "*Digital processing of speech signals*" Prentice-Hall, Inc., 1978

[9] W. Grifin and Jae S. Lim. "Signal estimation from modified short-time fourier transform" IEEE Trans., ASSP-32, NO 2, April 1984.

[10] James O. Pickles "*An Introduction to the Physiology of Hearing*" ACADEMIC PRESS 1988.

[11] Ray Meddis "Simulation of auditory-neural transduction: Further studies", JASA 83(3), March 1998 [12] Rey Meddis "Simulation of mechanical to neural transduction in the auditory receptor", JASA 79(3), March 1986.

[13] S. Odaira(1991) " *Simulation of speech processing in auditory nerves.*", Master Thesis, Shizuoka University, Japan (In Japanese).

[14] B. Delgutte "Representation of speech like sounds in the discharge patterns of auditory-nerve fibers", JASA, 68(3), (1980.9)

[15] Habibzadeh V. Houshang, S. Kitazawa "*Reverse* processing of the auditory-neural transducion for auditory model inversion ", Proceeding of ASJ, 2-8-16 pp-377, March 1997.