J. Acoust. Soc. Jpn. (E) 21, 4 (2000)

PAPER

An approach to auditory neural transduction reverse model

Houshang Habibzadeh Vaneghi and Shigeyoshi Kitazawa

Graduate School of Science and Engineering, Shizuoka University, 3-5-1, Johoku, Hamamatsu, 432-8011 Japan

(*Received 28 October 1999*)

In this paper we have addressed the problem of resynthesizing stimulus signal from adapted auditory neural firing pattern. The major issues discussed are: new method of nonlinearity inversion and effects of the stimulus signal's properties in forward and reverse outputs. A simple and efficient inner hair cell (IHC) inversion method based on Meddis IHC simulation has been constructed. With this inversion method and in response to tone-bursts of increasing frequencies and intensities, it was possible to reverse-process the nonlinearity of the auditory system and regenerate the estimate of the stimulus signal. Estimated signal showed good recovery of the information such as amplitude, frequency and phase, even in frequencies above 3 kHz. However, at intensities higher than 65 dB, amplitude recovery was not satisfactory enough. We conclude that, in reverse auditory simulations, our inversion method recovers important information pertaining to the identity of the original stimulus signal and could be employed as an IHC output monitoring or evaluation system.

Keywords: Inner hair cell, Reverse auditory model, Synaptic model, AGC, Auditory neural firings

PACS number: 43. 64. Bt, 43. 71. Cq

1. INTRODUCTION

In recent years, a number of sophisticated computational and experimental auditory fiber activity studies have been presented that aim to explain the particular nonlinearities that occur at the junction between the inner hair cells (IHC) and auditory-nerve fibers, the point of neuromechanical transduction.[1-3] So far, forward auditory models have been used as a powerful tool for investigating the neural mechanisms underlying speech perception. However, unlike other fields of science and engineering, usability and application domain of its reverse models have not been explored yet. Particularly, it is evident that despite a number of forward IHC models,[4-8] there has not been any remarkable study on IHC reverse simulations, their functionality, characteristics and possible applications.

In reverse applications, there are some transformations, like FFT, that can be inverted without loosing any information. However models containing nonlinear transformations such as many auditory and IHC models, are not 100% reversible. Then, a fundamental question is that, how a particular IHC inversion model will be useful in speech processing and what type of information will be recovered from inversion of the nonlinear transformations. To answer the above questions, first, we will explain about the possible application domain of the auditory inversion methods. Then we will investigate whether basic and essential information necessary for that application are present in the inverted signal.

In this paper, for IHC inversion purpose, we will rely on the Meddis IHC model[1,9] which is widely accepted as a realistic inner hair cell firing simulation. This model gives us the opportunity to setup a realistic and relatively simple model of human peripheral auditory system.[10] It is rich in parameters, which could be tuned manually or automatically.[11]

Due to the usage of the auditory or neural firing software/hardware models in speech recognition systems (as a pre-processor),[12] cochlear implant systems (as a hardware simulation),[13] and speech enhancement applications,[14,15] we need to investigate the outputs of the models to evaluate the condition and quality of the transferred signal.

For auditory model evaluation, we can reverse process the auditory model and resynthesis the estimate of the original stimulus signal.[10,14,15] Then the estimated signal could be checked with listening test. According to the idea of testing by resynthesis,[16] two different acoustic signals having the same auditory representation should sound identical to some extent. Therefore, degree of the perceptual difference between original and estimated signals is a good measure of the model's quality. Using this method, we may perceive original and estimated signals equivalent while they have different representations in some respect (*e.g.* different waveform).

The same argument and method can be applied for evaluation of the cochlear implant system (CIS) device. CIS itself is a prosthetic device that gets speech signal as input and generates electrically simulated auditory neural firing pattern in the output to provide limited speech comprehension. This device is being implanted as a replacement to peripheral auditory system in hearing impaired patients. Using the inversion method, electrically stimulated auditory perception of the CIS can be inverted by reverse processing of an auditory model. By regenerating the estimate of the original input signal, we are able to evaluate the device with test on normal listeners and avoid troublesome clinical preparations. Thus, many CIS experiments and tests could be done just in a laboratory and with computer simulations.

Another approach to the auditory model inversion is to use it in pitch perception and sound separation systems.[13,14] This method uses the temporal periodicity information and sound grouping capability of the auditory system for recovering of information in the two successive inversion stages. From a correlogram (representation of sound as a three dimensional function of time, frequency and periodicity) to a cochleagram, and then from a cochleagram to a waveform.

In the rest of the paper, to verify the richness of the estimated signal from IHC inversion, recovery rate of the amplitude, frequency and phase will be investigated.

It is to mention that, some primitive auditory inversion ideas are seen in the literature.[16,17] The only satisfactory auditory and adaptation inversion study prior to our work carried out by D. Naar[14] and M. Slaney[15] which was based on Lyon's cochlear model.[7] In the next section we will give a brief explanation about Lyon's model adaptation and its weakness against noise and signal level.

2. INVERSION OF THE LYON'S MODEL

In the Lyon's cochlear model inversion, the cochlear output's firing pattern was inverted by undoing the automatic gain control (AGC), finding the missing portions of the waveform that were removed by the detector, and combining the channels of the filter bank to create a waveform that will generate the same firing pattern. This model used complicate and multi-stage AGC for adaptation simulation.[11] In the Lyon's model, output of the forward AGC, y, in response to input x and gain G is defined to be:

$$y=Gx$$

The estimated X_{es} from inversion processing is related to y by:

$$X_{es}=(y+N)/G_{es} \quad (N \text{ is noise})$$

Substituting the relation for y results in

$$X_{es}=(Gx+N)/G_{es}=(G/G_{es})x+N/G_{es}$$

The first coefficient in the above equation causes the resulting estimate to be scaled by (G/G_{es}) if the gain and the gain estimate are not equal. If G is small, then an error in G_{es} translates into a large multiplicative and additive noise error in inversion. With very large signals, the AGC state is pushed close to one and the gain hovers near zero. Small amounts of noise sent back through the AGC state estimator translate into large changes in gain when the AGC is inverted.

In an attempt to overcome above-mentioned problems concerning AGC adaptation, a new nonlinearity inversion algorithm based on Meddis IHC model will be defined.

3. INVERSION OF THE MEDDIS IHC MODEL

In this section, the IHC model will be simulated in forward with intensity and frequency varying tone bursts. Then, the model outputs will be used in inversion process to estimate the stimulus signal with recovering the effect of the nonlinear transformations. Finally, degree of recovery from these transformations and synchronization effect will be investigated.

3.1 Meddis IHC Model

Meddis IHC model, which has been described in detail by Meddis,[1] can be viewed in terms of the production, movement, and dissipation of the transmitter substances in the region of the inner hair cell and the auditory nerve fibers. Figure 1 and differential equations (1)-(4) define the model's structure.

$$dq/dt = y(M - q(t)) + xw(t) - k(t)q(t) \qquad (1)$$

$$dc/dt = k(t)q(t) - lc(t) - rc(t) \qquad (2)$$

$$dw/dt = rc(t) - xw(t) \qquad (3)$$

$$k(t) = gdt(S(t)+A)/(S(t)+A+B), \quad S(t)+A>0$$
$$k(t)=0, \quad S(t)+A<0 \qquad (4)$$

The nonlinear $k(t)$ function is intended to reflect the permeability of the membrane. Transmitter substance is assumed to be stored in a free pool q near the synaptic junction and to be released across the membrane into the cleft c. The permeability factor $k(t)$ of the membrane is a nonlinear function of the instantaneous amplitude of the stimulus signal $S(t)$. Symbols g, A and B are permeability parameters. Factors r, l, x and y are refactory, loose, reprocessing and replenish coefficients respectively. Firing-rate factor, h, is a scale factor and M is transmitter quantity.[9]

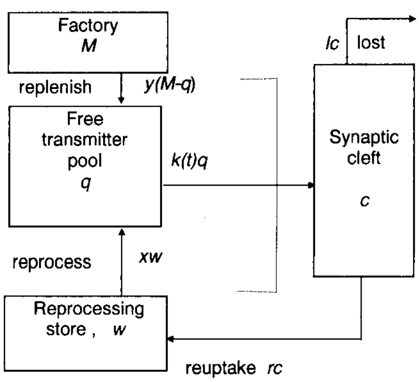


Fig. 1 Meddis inner hair cell model B.[1]

3.2 Meddis IHC Inversion Algorithm

Since no analytic solution exists for the Meddis original differential equations except for the period of silences,[1] those original equations and equations (5)-(8) which are derived from them are evaluated numerically and iteratively using small time intervals dt.

Getting the probability of firing rates or synaptic cleft data as input and launching the following algorithm, we are able to estimate input acoustic pattern (S_{est}) to the model. Estimated signal would be perceptually equivalent to the original input acoustic pattern. Table 1 shows the input and output relations between Meddis IHC forward and reverse models.

The following algorithm defines the Meddis IHC inversion process for a single fiber:

1. Initialize the reservoir contents to $w=0.1628$ and $q=0.1187$.
2. Get cleft contents $c(t)$ or neural firing rate data for sampling period dt.
3. Use the Eq. (5) (derived from Eq. (2)) and calculate the membrane's permeability function $k(t)$

$$k(t) = [c(t+dt) + c(t)(rdt + ldt - 1)]/q(t) \qquad (5)$$

4. Use Eq. (6) and (7) (derived from Eq. (1) and (2) respectively), adjust the contents of the reservoirs $q(t)$ (transmitter pool) and $w(t)$ (reuptake pool)

$$q(t+dt) = q(t) - k(t)q(t) + (xw(t) + y(M - q(t)))dt \qquad (6)$$

$$w(t+dt) = w(t) + (rc(t) - xw(t))dt \qquad (7)$$

Table 1 Input/Output relations between Meddis IHC forward and reverse models.

Model	Input	Output
Forward model	$S(t)$	$c(t)$
Reverse model	$c(t)$	$S_{est}(t)$

5. Use Eq. (8)(derived from Eq. (4)) and $k(t)$ to estimate the instantaneous stimulus data, $S_{HWR}(t)$, for dt epoch.

$$S_{HWR}(t)=\frac{gAdt-k(t)(A+B)}{k(t)-gdt} \quad (8)$$

6. Output the estimated signal $S_{HWR}(t)$, that is the half wave rectified (HWR) version of the estimated signal S_{est}.

The reservoirs initial values in reverse model have been defined based on reservoir contents at steady state of the forward model. However, using other positive and less than one initial values will converge to the same results over longer simulation time. Following experiments verify the capability of the inversion algorithm in response to various stimulus signal.

3.3 Simulations and Experiments

In this subsection, we outline the result of reverse and forward simulations with different input stimulus. Experiments and calculations have been done with MATLAB^{tm} and Mathematica^{tm} programming under Windows^{tm} operating system. All parameters (Table 2) and stimulus signals employed in our experiments were in compliance with Ray Meddis' last paper.[9]

Forward simulations have been carried out to re-investigate the characteristics of the Meddis IHC model, and use its outputs as inputs for the inverse model.

3.3.1 Processing with tone bursts of increasing intensities

3.3.1.1 Forward processing

In the first forward simulation which is the re-examination of the Meddis' experiment,[1] the stimulating signal was combination of the increasing intensities of a 1-kHz, 300 ms sinusoidal tone bursts with 250 ms intervening silences. Tone intensities varied from 40 dB to 80 dB with 5 dB steps ($S(t)$ is normalized such that $S^2(t)=1$ corresponds to a sound level of approximately 30 dB re μPa). At this and all other experiments, the sampling frequency (SF) was 20 kHz, unless otherwise stated. Figure 2 shows the response of the Meddis model to the mentioned signal. Note that with intensity increase, onset rate is increasing while steady state rate remains almost unchanged for different intensities (Fig. 2(c)). Model shows very good level of adaptation in different intensities and silence period that agrees with physiological findings.[2]

3.3.1.2 Reverse processing

In this experiment, neural firing or cleft contents data, captured from forward experiment explained in subsection 3.3.1.1 (Fig. 2(a)), have been used as an input to Meddis IHC inversion system. Using inversion algorithm in time intervals dt, estimate of the input tone burst to the Meddis model has been calculated. Figure 3 shows the estimated signal and its degree of recovery from onset and steady state firings. The steady-state function represents the firing rate after adaptation to the stimulus tone. The onset function represents the firing rate in the first millisecond after tone onset.[18] Figure 3(b) shows that except over 65 dB, recovery from onset firing is satisfactory.

To reduce the error, we tried shorter time intervals up to 0.025 ms for numerical evaluations of the differential equations. This did not have noticeable effect on results. However, steady state level

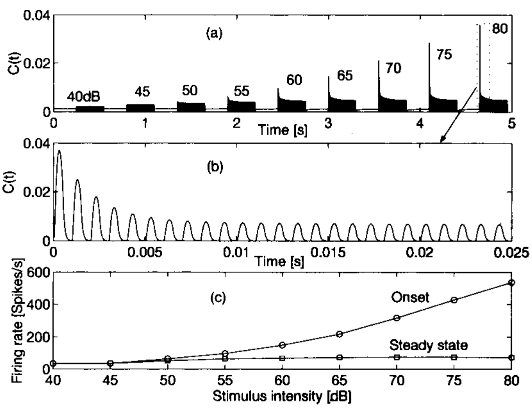


Fig. 2 Response of the Meddis model to 1-kHz 300 ms tone bursts with 5 dB steps and 250 ms intervening silences. (a) Synaptic cleft contents. (b) First 25 ms of the 80 dB, 1-kHz tone burst. (c) Onset and steady state firing rates at different intensities.

Table 2 Meddis IHC models' parameters set.

A	B	y	g	l	r	x	M	h	dt
5	300	5.05	2,000	2,500	6,580	66.31	1	50,000	0.05 ms

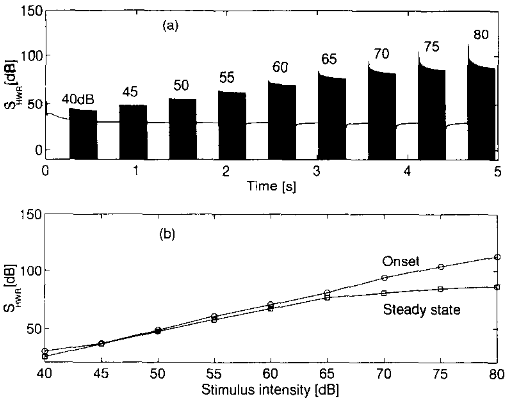


Fig. 3 Estimated 1-kHz 300 ms tone bursts with 5 dB steps and 250 ms intervening silences. (a) Estimated amplitude of the S_{HWR} signal using inversion process. (b) Recovery of S_{HWR} from onset and steady state firings at different intensities.

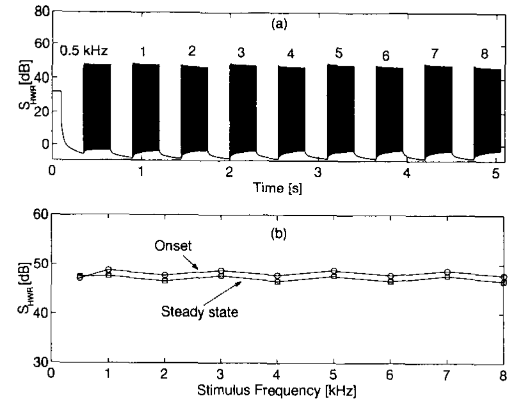


Fig. 5 Estimated tone bursts at frequency range from 0.5 kHz to 8 kHz with about 1-kHz steps. (a) Estimated 50 dB S_{HWR} signal with increasing frequencies. (b) Recovery of S_{HWR} from onset and steady state firings at different frequencies.

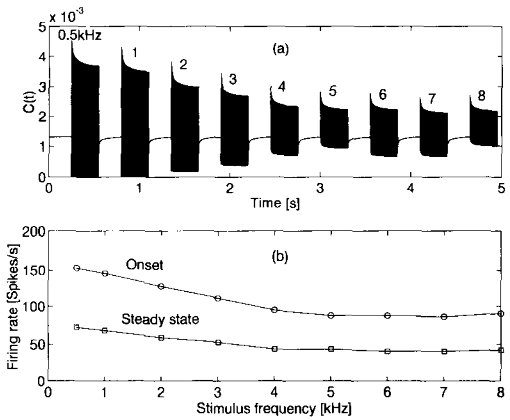


Fig. 4 Response of the Meddis IHC model to 50 dB tone bursts at frequency range of 0.5-8 kHz with 1-kHz steps. (a) Synaptic cleft contents. (b) Onset and steady state firing rates.

recovery was acceptable for whole intensity ranges with slight mismatching in higher intensities (Fig. 3(b)).

3.3.2 Processing with tone bursts of increasing frequencies

3.3.2.1 Forward processing

Figure 4 shows the amount of the synaptic cleft data, resulting from a series of tones with different frequencies. Signal intensity was set to 50 dB, and frequency range was from 0.5 kHz to 8 kHz with about 1-kHz step. Here, in contrast with intensity variations, with increase of the frequencies, onset rate, steady state rate and periodic part of the signal are decreasing while DC part of the $c(t)$ is increasing (Fig. 4(a)). This behavior completely agrees with the earlier physiological studies.[19]

3.3.2.2 Reverse processing

Synaptic cleft contents data, captured from forward experiment in subsection 3.3.2.1 (Fig. 4(a)), have been used as an input to Meddis IHC inversion system. Figure 5 (a) shows the output from inversion process. As it can be seen, inversion from tone bursts of increasing frequencies from 0.5 kHz to 8 kHz with 50 dB fixed intensity, provide us with acceptable estimate of the original signal. However, due to very compressed time scale in Fig. 5(a), estimated tone frequencies and shape of the cycles are not visible. Figure 5(b) shows that recovery from onset and steady state firing is near perfect.

3.3.3 Phase locking in forward and reverse simulations

It is clear that the ability of the model's excitation function to reflect the fine structure of the stimulus is limited by the rate at which the transmitter could be cleared from the cleft. This clearance is affected by two factors: dissipation and reuptake into the cell. When these two are slow relative to stimulus frequency, phase locking will be less evident.[18]

The forward synchronization coefficients, S_c, have been calculated using the relation (9):

$$S_c(\%) = \{(\text{Number of spikes in half cycle}) / (\text{Number of spikes in full cycle})\} \cdot 100 \quad (9)$$

Note that, the chosen half cycle should be the most populated one in that period.

To test the phase locking of the model in forward and reverse, 60 dB sinusoidal stimuli of 1,2,3,4 and 5 kHz were used with sampling frequency of 100 kHz. The cleft contents were averaged over one whole cycle of the signal then multiplied by a rate factor in order to estimate the approximate firing rate in spikes per second for that cycle. Figure 6(a) shows forward synchronization coefficient as a function of stimulus frequency. As it is seen, the average synchronization rate is about 65% and coefficients decline in strength between 1 and 5 kHz. This behavior agrees with the earlier empirical data.[20]

The reverse model's synchronization coefficients, RS_c, are based on each period of S_{HWR} and represents mean of the populated half of the cycle as a percentage of mean of its full cycle. Relation (10) has been used to calculate the reverse synchronization coefficients for estimated signal. This relation is extracted from relation (9) by eliminating the rate factor h from nominator and denominator.

$$RS_c(\%) = \{(\text{mean of the populated half cycle}) / (\text{mean of the full cycle})\} \cdot 100 \quad (10)$$

Figure 6(b) shows the synchronization coefficient of the estimated signal for different frequencies.

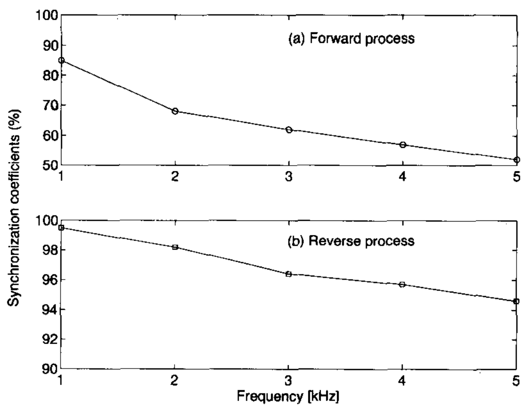


Fig. 6 Synchronization coefficients as a function of stimulus frequency. (a) Meddis IHC model's behavior (After Ray Meddis[18]). (b) Reverse IHC model's behavior.

From Fig. 6(b) we can see that phase information recovery rate is very high even in higher frequencies. However, phase recovery slightly decreases when stimulus frequency increases.

4. DISCUSSION

In auditory models, which are nonlinear in nature, auditory stimulus signal has some information like characteristic periodicity that remains undisturbed by most of its nonlinear transformations. While some information such as bandwidth, amplitude and phase characteristics of signal are changing in some degree.[15,21] Number of disturbed information and degree of distortion of the signal depends on the structure of the nonlinearities and nature of the stimulus signal (*e.g.* its frequency and intensity).

4.1 Recovered Information in Higher Frequencies

According to the outputs of the Meddis IHC inversion experiments, estimated signal at 50 dB from tone-bursts of increasing frequencies (Fig. 5), has perfect timing, amplitude and periodicity information in all working frequency ranges.

Earlier studies on analysis of the responses of the auditory neurons against a frequency varying tone-burst, relied only on data computed from PST histograms for representation of firing rate in high frequencies. Thus, due to low firing rate in frequencies higher than 1.5 kHz, it was concluded that the conveyed information in this region is very poor[22] (*e.g.* poor localized synchronized rate). However, our inversion results show that, even in higher frequencies there are enough conveyed information which enables us to regenerate the stimulus signal with acceptable recovery of the information such as, amplitude, timing and phase.

4.2 Onset Firing Recovery in Higher Intensities

In case of inverting stimulus signal containing increasing intensities, due to overshooting phenomena at intensities above 65 dB (Fig. 2), onset rate firing recovery was not satisfactory. This problem can be viewed in two ways. First, it is related to physiological behavior of auditory system in response to abnormal high intensity stimulus, which starts with sudden and very high firing rate, and then declines to a steady state level. Second, it is related to the inversion system itself, which over estimates the data and fails to handle very high onset firing at intensities higher than 65 dB. This distortion is the auditory system and regenerate the estimate of the stimulus signal, which recovered most of the information needed for auditory or CIS output monitoring or evaluation.

ACKNOWLEDGMENTS

The authors are grateful to Prof. R. Meddis of University of Essex for his helpful comments on an early version of this manuscript; to Dr. M. Slaney of Intervals Research Corp. for his useful advises and also to Dr. T. Kitamura for many constructive discussions.

REFERENCES

1) R. Meddis, "Simulation of mechanical to neural transduction in the auditory receptor," J. Acoust. Soc. Am. **79**, 702–711 (1986).
2) L. A. Westerman and R. L Smith, "Rapid and short term adaptation in auditory-nerve responses," Hear. Res. **15**, 249-260 (1986).
3) M. R. Schroeder and J. L. Hall "Model for mechanical to neural transduction in the auditory receptor," J. Acoust. Soc. Am. **55**, 1055-1060 (1974).
4) Y. Oono and Y. Sujaku, "A model for automatic gain control observed in the firings of primary auditory neurons," Trans. Inst. Electron. Commun. Eng. Jpn. **58**, 352–358 (1975).
5) S. Seneff, "A joint synchrony/mean-rate model of auditory processing," J. Phonet. **16**, 55-76 (1988).
6) T. Furukawa and S. Matsuura, "Adaptive rundown of excitatory post-synaptic potentials at synapses between hair cells and eighth nerve fibers in goldfish," J. Physiol. **276**, 193-209 (1978).
7) R. F. Lyon, "A computational model of filtering, detection, and compression in the cochlea," IEEE Trans. Acoust. Speech Signal Process. **12**, 1282-1285 (1982).
8) L. A. Westerman and R. L. Smith, "A diffusion model of the transient response of the cochlear inner hair cell synapse," J. Acoust. Soc. Am. **83**, 2266–2276 (1988).
9) R. Meddis, M. J. Hewitt, and T. M. Shackleton, "Implementation details of a computational model of the IHC/AN synapse," J. Acoust. Soc. Am. **87**, 1813–1816 (1990).
10) H. H. Vaneghi and S. Kitazawa, "Meddis IHC based multi-channel cochlear model setup and inversion using three types of auditory filter banks," Proc. 14th Int. Congr. Phonetic Sciences (ICPhS99), Vol. 1, 211–214 (1999).
11) Chen P. Wong and M. J. Pont, "Automatic selection of parameters for a computer simulation of extracellular auditory nerve fiber activity," Proc. ESCA Workshop on the Auditory Basis of Speech Perception, 61–64 (1996).
12) P. Cosi, "Auditory modeling for speech analysis and recognition," in *Visual Representation of the Speech Signals*, M. Cooke, S. Beet, and M. Crawford, Eds. (J. Wiley & Sons Ltd., Chichester, 1993), pp. 205–212.
13) G. M. Clark, P. J. Blamey, A. M. Brown. P. A. Gusby, R. C. Dowell, B. K. -H. Franz, B. C. Pyman, R. K. Shepherd, Y. C. Tong, R. L. Webb, M. S. Hirshorn, J. Kuzma, D. J. Mecklenburg, D. K. Money, J. F. Patrick, and P. M. Seligman, "*The University of Melbourne Nucleus Multi-electrode Cochlear Implant*," *Advances in oto-rhino-Laryngology*, Vol. 38 (KARGER AG, Basel, Switzerland, 1987), p. 129.
14) D. Naar, "Sound resynthesis from a correlogram," San Jose Univ., Dept. of Elect. Eng.,Tech. Rep. No. 3 (1993).
15) M. Slaney, D. Naar, R. F. Lyon, "Auditory model inversion for sound separation," Proc. ICASSP 94, Vol. 2, 77–80 (1994).
16) R. W. Hukin and R. I. Damper, "Testing an auditory model by resynthesis," Eur. Conf. Speech Communication and Technology, Vol. 1, 243–246 (1989).
17) X. Yang, K. Wang, and S. Shamma, "Auditory representation of acoustic signals," IEEE Trans. Inf. Theory **38**, 824-839 (1992).
18) R. Meddis, "Simulation of auditory-neural transduction: Further studies," J. Acoust. Soc. Am. **83**, 1056–1063 (1998).
19) J. O. Pickles, *An Introduction to the Physiology of Hearing*, 2nd Ed. (Academic Press, London, 1988), p. 60.
20) J. E. Rose, J. F. Brugge, D. J. Anderson, and J. E. Hind, "Phase locked response to low frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. **30**, 767-793 (1967).
21) M. Slaney and R. F. Lyon, "On the importance of time—A temporal representation of sound," in *Visual Representation of Speech Signals*, M. Cooke, S. Beet, and M. Crawford, Eds. (J. Wiley & Sons Ltd., Chichester, 1993), pp. 95–116.
22) E. D. Young and M. B. Sachs, "Representation of steady-state vowels in the temporal aspect of the discharge pattern of populations of auditory-nerve fibers," J. Acoust. Soc. Am. **66**, 1381–1403 (1979).
23) H. H. Vaneghi and S. Kitazawa, "Stimulus signal estimation from auditory neural transduction inverse processing," Proc. ESCA. Eurospeech97., ISSN-1018-4074, Vol. 5, 2795-2798 (1997).
24) H. H. Vaneghi and S. Kitazawa, "Reverse processing of the auditory-neural transduction for auditory model inversion," Proc. Spring Meet. Acoust. Soc. Jpn., ISSN 1340-3168, 2-8-16, 377–378 (1997).

source of a little bit added noise which appears in estimated signal at high intensities.

Estimated signal showed very good recovery of the steady state rate (Fig. 3(b)). Although here also, minor mismatching was found at intensities above 65 dB.

4.3 Effect of Sampling Frequency

To study the effect of the sampling frequency, particularly on onset firing recovery, we tried higher and lower sampling rates than 20 kHz. In 40 kHz sampling rate, there was not noticeable improvement in the results. Also, much lower sampling rate, *i.e.*, 5 kHz or lower, disables the rapid adaptation and causes the negative values appear in reservoirs and makes the system unstable.

4.4 Phase Recovery and Stimulus Frequency

Phase locking experiment on inverted signal showed that phase recovery is almost perfect in low frequencies. Recovery rate slightly decreases when stimulus frequency increases. However, synchronization coefficient is about 96% in average at frequency range of 1–5 kHz. Which shows better phase locking in compare to average synchronization rate of 65% in forward simulation.

4.5 Remaining Inversion Processes

After calculating S_{HWR} from single-fiber IHC inversion process, for a full and multi-channel auditory model inversion, the process will be completed with undoing the effects of the auditory filter and half-wave rectifier on S_{HWR} to get S_{est} for each channel. Final output as summation of all channels, will be the estimate of the original input signal to the peripheral auditory system.

4.6 Effect of the Un-recovered Information on the Estimated Signal

As stated before, in this paper we do not claim that auditory output is completely reversible. That is, because of some eliminated and unrecoverable information by adaptation operation, IHC's rectifying behavior and filtering property of the basilar membrane, estimated signal can not be equal to the original signal. However, in a human auditory system, the eliminated or reduced data in forward processing is believed to be an irrelevant or out of operating-range information. This irrelevance may be interpreted in the following way. In a perfect inversion of the auditory output with partial recovery of the reduced information, if the estimated signal were reprocessed on the same auditory system, it would produce the same firing pattern as original signal.

4.7 Informal Perceptual Evaluation of the Estimated Speech Signal

Informal perceptual tests on the estimated speech signal from our previous multi-channel IHC based auditory model inversion,[10,23,24] provided us with near perfect representation of the original stimulus signal. Although, there was a little bit added noise resulted from inversion of HWR and overshooting problem stated in subsection 3.3.1.2, good quality of the estimated signal can be interpreted as a good performance of the Meddis IHC based reverse model. Nevertheless, authors believe that to be more precise on conclusion, perceptual test of the original and estimated signal should be carried out in more controlled environment and with more listeners.

5. CONCLUSIONS

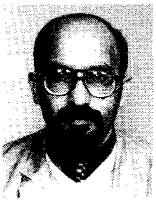
An inner hair cell inversion method based on Meddis IHC model has been defined. Using this efficient inversion method, we were able to achieve acceptable quality without employing complex multiplicative and multi-stage AGC of the earlier work.

The validity of the IHC inversion algorithm was confirmed with three sets of forward/reverse experiments. Inversion from tone bursts of increasing intensities was successful in recovering onset and steady state rate amplitude information up to 65 dB.

In reverse processing of the tone bursts of the increasing frequencies at fixed 50 dB, inversion method showed very good performance in recovering amplitude at frequency ranges up to 8 kHz. The method recovered amplitude even at frequencies higher than 3 kHz, where most of the information was reduced or eliminated in forward processing.

Phase recovery was quite satisfactory for 1–5 kHz range, even at frequencies higher than 1.5 kHz, where phase information was severely reduced in forward simulation. Nevertheless, recovery rate slightly decreased with increase of the stimulus frequency

With IHC inversion method, we could reverse process the nonlinearity of the Meddis IHC based



Houshang Habibzadeh Vaneghi was born in Sarab, Iran on 26 May 1959. He received Bachelor of Computer Engineering from Tehran University in 1986 and Master of Computer Science and Engineering from Shizuoka University in 1997. He is currently a Ph.D. candidate in Shizuoka University and expected to graduate in 2000. His research interests are forward and reverse auditory modeling and their applications in speech processing and Cochlear Implant Systems. He is the student member of ASJ, ASA, IEEE and its Computer and Signal processing societies.

Shigeyoshi Kitazawa was born in Wakayama, Japan on March 11 1949. He received B. E. E., M. E. E. and Doctor of Engineering degrees all from Kyoto University in 1971, 1973, and 1976, respectively. After working as a research associate of Kyoto University, he was an associate professor of Shizuoka University in 1985, and a professor in 1994. His research interests are feature extraction and discriminant analysis of phonemes, auditory model, and speech processing for cochlear implant.