# Using Values of the Human Cochlea in the Macro and Micro Mechanical Model for Automatic Speech Recognition

José Luis Oropeza Rodríguez and Sergio Suárez Guerra

Computing Research Center, National Polytechnic Institute,
Juan de Dios Batiz s/n, P.O. 07038, Mexico
{joropeza,ssuarez}@cic.ipn.mx

**Abstract.** Recently the parametric representation using cochlea behavior has been used in different studies related with Automatic Speech Recognition (ASR). That is because this hearing organ in mammalians is the most important element used to make a transduction of the sound pressure that is received by the outer ear. This paper shows how the macro and micro mechanical model is used in ASR tasks. The values that Neely, Elliot and Ku founded in their works, related with the macro and micro mechanical model such as Neely were used to set the central frequencies of a bank filter to obtain parameters from the speech in a similar form as MFCC (Mel Frequency Cepstrum Coefficients) has been constructed.

An approach that considers a new form to distribute the bank filter in our parametric representation is proposed. Then this distribution of the bank filter to have a different representation of the speech in frequency domain compared with MFCC is applied. The response of these three values mentioned above into macro and micro mechanical model to create the central frequencies of the bank filter were used, then the Mel scale function substituted by a representation based in the cochlear response based on the Neely model. This model was used with a set of different parameters of the cochlea, used by Nelly, Elliot and Ku in their works, such as mass, damping and stiffness; among others. A performance of 98 to 100% was reached for a task that uses Spanish isolated digits pronounced by 5 different speakers. Corpus SUSAS with neutral sound records with some advantages in comparison with MFCC was applied.

**Keywords:** Speech recognition, cochlea, place theory and bank filter.

## 1 Introduction

For a long time Automatic Speech Recognition Systems (ASRs) have used parameters related with Cepstrum and Homomorphic Analysis of Speech [1], Linear Prediction Coefficient (LPC) [2], Mel Frequency Cepstrum Coefficients (MFCC) [3], and Perceptual Linear Prediction (PLP) [4], these last two being the most important. In each of these representations, the principal objective is to have a representation to compress speech data without irrelevant information not pertinent to the phonetic analysis of the data and to enhance aspects of the signal that contribute significantly to the detection of phonetic differences. MFCC and PLP coefficients employ Mel and Bark

A. Gelbukh et al. (Eds.): MICAI 2014, Part II, LNAI 8857, pp. 242–251, 2014.

scales respectively. These consider perceptual aspects to obtain a set of coefficients that represent the speech signal.

On the other hand, the most important organ in human hearing is the cochlea and various physiological models have been proposed [5] and [6]. Recently works related with the application of the cochlea behavior in ASR systems can be found, that is because in recent years the researchers have emphasized "human engineering", that is, to adopt the processing strategies of the human auditory perception. The application of such a human perceptual feature may improve ASR performance which has been established in literature [7][8][9][10][11][12]. In [12] an extraordinarily precise auditory model was used extracting the excitation dependent shapes of the delay trajectories and then a set of features were used without any other spectral information to carry out speech recognition task under different noise conditions on the TIMIT database. However, average recognition rates do not reach that of the MFCC features (except for very low noise SNRs), but the system behaves very stable under different noise conditions. In [11] they proposed a feature extraction method for ASR based on the differential processing strategy of the AVCN, PVCN and the DCN of the nucleus cochlear. The method utilized a zero-crossing with peak amplitudes (ZCPA) auditory model as synchrony detector to discriminate the low frequency formants. They used HMM recognition using isolated digits that showed better recognition rates in clean and non- stationary noise conditions than the existing auditory model. In [10] they employed a counterpart of the next physiological processing step in comparison with frequency decomposition and compression of amplitudes concepts. A simplified model of short-term adaptation was incorporated into MFCC feature extraction. They compared the proposal mentioned above with that structurally related to RASTA, CMS and Wiener filtering which performs well in combination with Wiener filtering. Compared with the structurally related RASTA, the adaptation model provides superior performance on AURORA 2, and, if Wiener filtering is used prior to both approaches, on AURORA 3 as well.

## 2 Characteristics and Generalities

The cochlea is a long, narrow, fluid-filled tunnel which spirals through the temporal bone. This tunnel is divided along its length by a cochlear partition into an upper compartment called scala vestibuli (SV) and lower compartment called scala timpani (ST). At the apex of the cochlea, SV and ST are connected to each other by the helicotrema [13]. A set of models to represent the operation of the cochlea has been proposed [14][15][16][17]; among others. In mammals, vibrations of the stapes set up a wave with a particular shape on the basilar membrane. The amplitude envelope of the wave is first increasing and then decreasing, and the position at the peak of the envelope is dependent on the frequency of the stimulus [18]. The amplitude of the envelope is a two-dimensional function of distance from the stapes and frequency of stimulation. The curve shown in Fig. 1 is a cross-section of the function for fixed frequency. If low frequencies excite the cochlea, the envelope is nearest to the apex, but if high frequencies excite it the envelope is nearest to the base.

This paper proposes an equation extracted from the fluid mechanical model to find a relationship between these frequencies and the place of the excitation into the cochlea. With that value a new distribution of the bank filter to extract parameters for ASR tasks is proposed.

In the micromechanical the anatomical structure of a radial cross-section (RCS) of the cochlear partition (CP) is illustrated in the following figure 2. In the model, the basilar membrane (BM) and tectorial membrane (TM) are each represented as a lumped mass with both stiffness and damping in their attachment to the surrounding bone. When the cochlea determines the frequency of the incoming signal from the place on the basilar membrane of maximum amplitude, the organ of Corti is excited, in conjunction with the movement of tectorial membrane; the inner and outer hair cells are excited obtaining an electrical pulse that travels by auditory nerve.

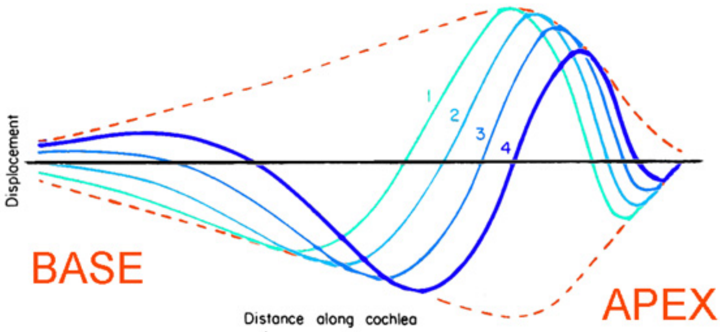
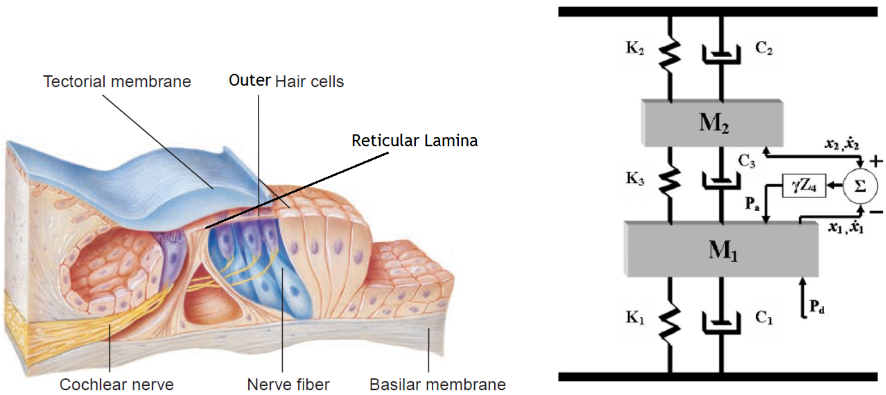



**Fig. 1.** Wave displacement inside cochlea




(a) (b)

**Fig. 2.** Anatomical structure of the cochlear partition (a). The outer hair cells, micro mechanical representation (b).

Now the modeling cochlear will be divided in two ways of study. The first is the hydrodynamic movement that produced a movement on the basilar membrane and the second is the movement of the outer hair cells. This is named as the model of Macro and Micro Mechanical Cochlear [17]. The equations that describe the Macro Mechanical Cochlear are [17]:

$$\frac{d^2}{dx^2} P_d(x) = \frac{2\rho}{H} \varepsilon_p \ddot{(x)}, \tag{1}$$

$$\frac{d}{dx} P_d(0) = 2\rho \ddot{\varepsilon_s}, \tag{2}$$

$$\frac{d}{dx} P_d(L) = 2\rho \ddot{\varepsilon_h}, \tag{3}$$

The equations (1), (2) and (3) were solved by finite difference, using central differences for (1), forward differences for the (2) and backward difference for (3), generating a tri-diagonal Matrix system[16] which we solved using the Thomas algorithm. It represents the Micro mechanical, because it uses the organ of Corti values.

$$\begin{bmatrix}
\left(\frac{2\rho i\omega}{Z_m} - \frac{1}{\Delta}\right) & \frac{1}{\Delta} & \cdots & 0 & 0 & & 0 & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & & & & & & & & & 0 \\
0 & 0 & \cdots & \frac{1}{\Delta^2} & -\left(\frac{2}{\Delta^2} + \frac{2\rho i\omega}{HZ_p(X_n)}\right) & & \frac{1}{\Delta^2} & \cdots & 0 & 0 \\
0 & \vdots & & \vdots & \vdots & & \cdots & & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & & & & & \vdots \\
0 & & & & & & \cdots & \ddots & \ddots & 0 \\
0 & 0 & \cdots & 0 & 0 & & 0 & \cdots & \frac{1}{\Delta} & -\left(\frac{1}{\Delta} - \frac{2\rho i\omega}{C_h}\right)
\end{bmatrix}
\begin{bmatrix} P_d(X_1) \\ P_d(X_2) \\ \vdots \\ P_d(X_{n-1}) \\ P_d(X_n) \\ P_d(X_{n+1}) \\ \vdots \\ P_d(X_{N-1}) \\ P_d(X_N) \end{bmatrix}
=
\begin{bmatrix} \left(\frac{2\rho i\omega A_m}{Z_m G_m A_s}\right) P_e \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \tag{4}$$

The solution for $P_d$ obtains the maximum amplitude on the basilar membrane shown in Figure 3. For these experiments the cochlear distance pattern is obtained manually. As can be seen, to solve equation 4 a set of variables related with the physiology of the cochlea is needed and some of these variables are described in table 1. These values are immersed into $Z_p$ and $Z_m$; for example in [17].

Figures 3, and 4; show the behavior of the basilar membrane with the values shown in table 1. As is seen, before 300 Hz the behavior of the micro and macro mechanical model is not adequate, independently of the parameters used. This result is a consequence of the characteristics of the model proposed by [17]. Proposing our analysis from this frequency to 4.5 KHz was decided. Also, the response obtained has a behavior logarithmic. This is an important indication because the Mel function is related with a similar mathematical function.

**Table 1.** Values used in equation

| *Parameter* | *Neely & Kim 1986 (cgs)* | *Ku (human cochlea, 2008)IS* | *Elliot (2007)IS* |
|---|---|---|---|
| $k_1(x)$ | $1.1 * 10^9 e^{-4x}$ | $1.65 * 10^8 e^{-2.79(x+0.00373)}$ | $4.95 * 10^8 e^{-3.2(x+0.00375)}$ |
| $c_1(x)$ | $20 + 1500e^{-2x}$ | $0.9 + 999e^{-1.53(x+0.00373)}$ | $0.1 + 1970e^{-1.79(x+0.00375)}$ |
| $m_1(x)$ | $3 * 10^{-3}$ | $4.5 * 10^{-4}$ | $1.35 * 10^{-3}$ |
| $k_2(x)$ | $7 * 10^6 e^{-4.4x}$ | $1.05 * 10^6 e^{-3.07(x+0.00373)}$ | $3.15 * 10^6 e^{-3.52(x+0.00375)}$ |
| $c_2(x)$ | $10e^{-2.2x}$ | $3e^{-1.71(x+0.00373)}$ | $11.3e^{-1.76(x+0.00375)}$ |
| $m_2(x)$ | $0.5 * 10^{-3}$ | $0.72 * 10^{-4} + 0.28710^{-2}x$ | $2.3 * 10^{-4}$ |
| $k_3(x)$ | $1 * 10^7 e^{-4x}$ | $1.5 * 10^6 e^{-2.79(x+0.00373)}$ | $4.5 * 10^6 e^{-3.2(x+0.00375)}$ |
| $c_3(x)$ | $2e^{-0.8x}$ | $0.66e^{-0.593(x+0.00373)}$ | $2.25e^{-0.64(x+0.00375)}$ |
| $k_4(x)$ | $6.15 * 10^8 e^{-4x}$ | $9.23 * 10^7 e^{-2.79(x+0.00373)}$ | $2.84 * 10^8 e^{-3.2(x+0.00375)}$ |
| $c_4(x)$ | $1040e^{-2x}$ | $330e^{-1.44(x+0.00373)}$ | $965e^{-1.64(x+0.00375)}$ |
| gamma | 1 | 1 | 1 |
| g | 1 | 1 | 1 |
| b | 0.4 | 0.4 | 0.4 |
| L | 2.5 | 3.5 | 3.5 |
| H | 0.1 | 0.1 | 0.1 |
| $K_m$ | $2.1 * 10^5$ | $2.63 * 10^7$ | $2.63 * 10^7$ |
| $C_m$ | 400 | $2.8 * 10^3$ | $2.8 * 10^3$ |
| $M_m$ | $45 * 10^3$ | $2.96 * 10^{-3}$ | $2.96 * 10^{-3}$ |
| $C_h$ | 0.1 | 0.1 | 0.1 |
| $A_s$ | 0.01 | $3.2 * 10^{-1}$ | $3.2 * 10^{-1}$ |
| $A_m$ | 0.35 cm^2 | 0.429-0.55 cm^2 | 0.429-0.55 cm^2 |
| Rho | 0.35 | 1 | 1 |
| N | 250 | 500 | 500 |
| Gm | 0.5 | 0.5 | 0.5 |

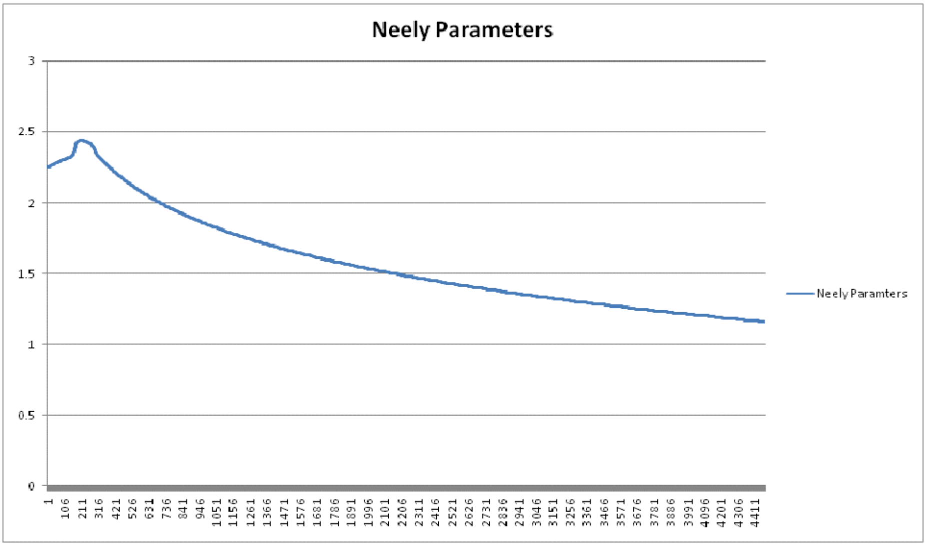



**Fig. 3.** Neely´s model using his parameters

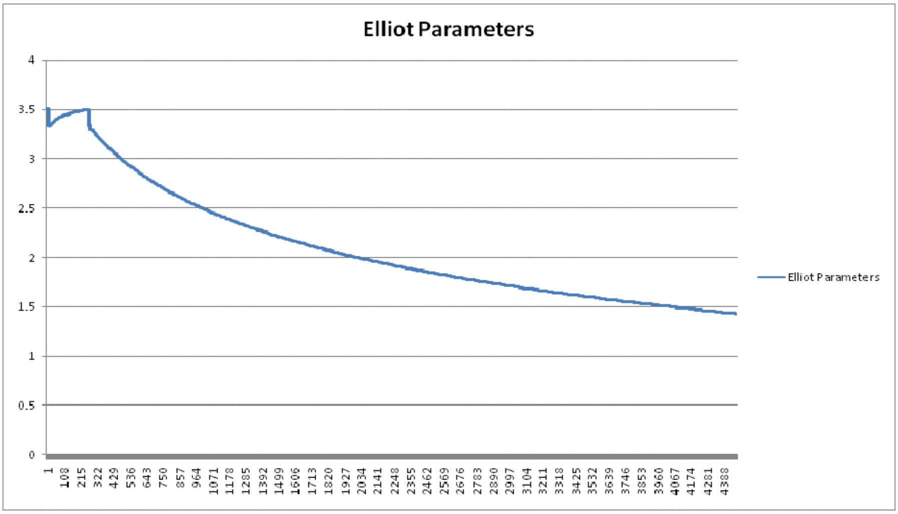



**Fig. 4.** Neely´s model using Elliot parameters

As mentioned above, the Neely model and later works have considered putting a number of these micro-mechanisms along the cochlea at the same distance between them. For that, this principle to establish the following relation between a minimal and maximal distance was used.

$$d(n) = dmax + \sum_{n=0}^{n=nint} n\frac{dmin - dmax}{nint + 1}, \tag{5}$$

Where $d_{min}$ and $d_{max}$ are obtained from Figure 3 and 4, considering that $F_{min}=300$ Hz and $F_{max}=4.5$ *KHz*. This paper proposed a space equidistant between different points to analyze the cochlea. After that, for each distance one specifically frequency of excitation to the Basilar Membrane was obtained.

## 3 Experiments and Results

From the last analysis a computational model to obtain the distance where the maximum displacement of the basilar membrane occurs to a specific excitation frequency of the system was developed, which depends of the physical characteristics of the basilar membrane. The following procedure describes the computational model of the cochlea using this propose [20]. It is important to mention that the maximum response of the pressure curve used in [19] was obtained.

1. Obtain speech signal, realize preprocessing (It includes pre-emphasis, segmentation, windowing and feature extraction), for each sentence.
2. In the feature extraction, the same procedure as MFCC was used but the filter bank is constructed following the next steps.

2.4 Determine the frequency related with these distances, this represents the center of the filter bank.

2.5 Construct filter bank with frequency center obtained from the analysis of the Neely model using values in table 1.

2.1 Take the minimal and maximal frequency where filter bank are going to be constructed.

2.2 Calculate maximal and minimal distance from the stapes of the cochlea, nearer to start implies high frequencies, farthest implies low frequencies.

2.3 Determine a set of distances equally spaced

3. Follow the same steps to obtain MFCC, multiply spectral representation from Fourier Transform with filter bank, calculate energy by bands using logarithm, and finally, apply discrete cosine transform.

4. Obtain a new set of coefficients for each speech signal.

5. Train the ASR and proceed with recognition task using the new parameters.

A database with 5 speakers that pronounced Spanish isolated digits, from 0 to 9 was applied as workbench that is "cero, uno, dos, tres, cuatro, cinco, seis, siete, ocho and nueve". LPC, MFCC, CLPC were used and our coefficients named EPCC (Earing Perception Cepstrum Coefficients) obtaining better percent correct recognition in some tasks using them in comparison with others representations mentioned above. HTK Hidden Markov Model Toolkit was used as training and recognition software; our new parameters were added into HSigp.c file, contained inside HTK http://htk.eng.cam.ac.uk, and were used in tasks of ASR employing HTK.

This first experimental used a database that contains only digits in the Spanish language and the characteristics of the samples were frequency sample 11025, 8 bits per sample, PCM coding, mono-stereo. The evaluation of the experiment proposed involved 5 people (3 men and 2 women) with 300 speech sentences to recognize for each one ( 100 for training task and 200 for recognition task). 1500 speech sentences extracted from 5 speakers individually were taken, and the Automatic Speech Recognition trained using Hidden Markov Models with 6 states (4 states with information and 2 dummies to connection with another chain). Also, 3 Gaussian Mixture for each state in the Markov chain were employed. The parameters extracted from the speech signal were 39 (13 MFCC, 13 delta and 13 energy coefficients) when using MFCC or our proposal, and used to train the Hidden Markov Model. Table 1 contains results obtained in percentages when using LPC, CLPC, MFCC and our parametric representation to train as parameters. Table 2 shows results using Delta and Acceleration coefficients. It is important to mention that HTK give us results in two forms: by sentence and by words http://htk.eng.cam.ac.uk. We show both for reasons of consistency. Table 3 contains results obtained in percentage when using LPC, CLPC and MFCC, DELTA, ACCELERATION AND THIRD DIFFERENTIAL.

**Table 2.** LPC, CLPC and MFCC coefficients

| ***SENTENCES*** | | | | ***WORDS*** | | | |
|---|---|---|---|---|---|---|---|
| PARAMETERS/# STATES | ***4*** | ***5*** | ***6*** | PARAMETERS/# STATES | ***4*** | ***5*** | ***6*** |
| LPC | 87.5 | 94 | 94 | LPC | 87.94 | 94.47 | 94.47 |
| CLPC | 90 | 97.5 | 98.5 | CLPC | 90.45 | 97.99 | 98.99 |
| MFCC | 97.5 | 97 | 99 | MFCC | 97.99 | 97.49 | 99.5 |
| EPCC KU | 98 | 99 | 99.5 | EPCC KU | 98.45 | 99.5 | 99.8 |
| EPCC ELLIOT | 98.5 | 98.5 | 99 | EPCC ELLIOT | 98.75 | 98.75 | 99.5 |
| EPCC NEELY | 98.7 | 99 | 99.5 | EPCC NEELY | 98.5 | 99.5 | 99.75 |

**Table 3.** LPC, CLPC, MFCC, DELTA AND ACCELERATION coefficient

| *SENTENCES* | | | | *WORDS* | | | |
|---|---|---|---|---|---|---|---|
| PARAMETERS/# STATES | ***4*** | ***5*** | ***6*** | PARAMETERS/# STATES | ***4*** | ***5*** | ***6*** |
| LPC | 79 | 90.5 | 91.5 | LPC | 79.4 | 99.4 | 91.96 |
| CLPC | 93 | 99 | 99 | CLPC | 93.47 | 99.5 | 99.5 |
| MFCC | 99 | 99 | 99 | MFCC | 99.5 | 99.5 | 99.5 |
| EPCC KU | 100 | 100 | 100 | EPCC KU | 100 | 100 | 100 |
| EPCC ELLIOT | 100 | 100 | 100 | EPCC ELLIOT | 100 | 100 | 100 |
| EPCC NEELY | 100 | 100 | 100 | EPCC NEELY | 100 | 100 | 100 |

**Table 4.** LPC, CLPC, MFCC AND DELTA, ACCELERATION, DELTA, AND THIRD DIFFERENTIAL coefficients

| *SENTENCES* | | | | *WORDS* | | | |
|---|---|---|---|---|---|---|---|
| PARAMETERS/# STATES | ***4*** | ***5*** | ***6*** | PARAMETERS/# STATES | ***4*** | ***5*** | ***6*** |
| LPC | 77 | 89.5 | 89 | LPC | 77.39 | 89.95 | 89.45 |
| CLPC | 89.5 | 99 | 99 | CLPC | 89.95 | 99.5 | 99.5 |
| MFCC | 98.5 | 99 | 99 | MFCC | 98.99 | 99.5 | 99.5 |
| EPCC KU | 100 | 100 | 100 | EPCC KU | 100 | 100 | 100 |
| EPCC ELLIOT | 100 | 100 | 100 | EPCC ELLIOT | 100 | 100 | 100 |
| EPCC NEELY | 100 | 100 | 100 | EPCC NEELY | 100 | 100 | 100 |

In the second experiment, a corpus elaborated by J. Hansen at the University of Colorado Boulder was used. He has constructed database SUSAS (Speech Under Simulated and Actual Stress) http://catalog.ldc.upenn.edu/LDC99S78.Only 9 speakers

**Table 5.** Results obtained using HTK, SUSAS Corpus and manual labeling

| | ***MFCC*** | | ***EPCC Using Neely values*** | | ***EPCC Using Ku values*** | | ***EPCC Using Elliot values*** | |
|---|---|---|---|---|---|---|---|---|
| | ***sen-tence*** | ***word*** | ***sen-tence*** | ***word*** | ***sen-tence*** | ***word*** | ***sen-tence*** | ***word*** |
| ***boston1*** | 91.84 | 92.06 | 90.61 | 90.87 | 90.2 | 90.48 | 89.39 | 89.68 |
| ***boston2*** | 95.51 | 95.63 | 93.47 | 93.65 | 93.47 | 93.65 | 93.06 | 93.25 |
| ***boston3*** | 96.73 | 96.83 | 93.88 | 94.05 | 95.92 | 96.03 | 96.33 | 96.43 |
| ***general1*** | 96.73 | 96.83 | 92.24 | 92.46 | 93.88 | 94.05 | 93.88 | 94.05 |
| ***general2*** | 94.29 | 94.44 | 90.61 | 90.87 | 90.61 | 90.87 | 89.39 | 89.68 |
| ***general3*** | 93.47 | 93.65 | 88.16 | 88.49 | 93.47 | 93.65 | 93.06 | 93.25 |
| ***nyc1*** | 91.84 | 92.06 | 91.84 | 91.67 | 87.35 | 87.3 | 96.33 | 96.43 |
| ***nyc2*** | 91.02 | 91.27 | 91.84 | 92.06 | 86.53 | 86.9 | 93.88 | 94.05 |
| ***nyc3*** | 95.92 | 96.03 | 92.65 | 92.86 | 90.61 | 90.87 | 89.39 | 89.68 |

with ages ranging from 22 to 76 were used and we applied normal corpus not under Stress sentences contained into corpus. The words were "brake, change, degree, destination, east, eight, eighty, enter, fifty, fix, freeze, gain, go, hello, help, histogram, hot, mark, nav, no, oh, on, out, point, six, south, stand, steer, strafe, ten, thirty, three, white, wide, & zero". A total of 4410 files of speech were processed. Finally, Tables 4 and 5 show results when using our proposal (Earing Perceptual Cepstrum Coefficients –EPCC-) the best representations used in the state of the art and in the last experiment versus MFCC in SUSAS corpus.

## 4 Conclusions and Future Works

This paper describes new parameters for ASRs tasks. They employ the functionality of the cochlea, the most important hearing organ of humans and mammalians. At this moment, the parameters used for the MFCC analysis have been demonstrated to be the most important parameters and the most used for this task. For many years a great diversity of models that attempt describing the functionality of the ear have been proposed and are implicit that this phenomenon has been used for ASR based on phenomenological models. Another alternative was employed using a physiological model based in macro and micro mechanical proposed by Neely. This article demonstrated that the Neely cochlea model can be used to obtain speech signal parameters for Automatic Speech Recognition. In conclusion, the cochlea behavior can be used to obtain these parameters and the results are adequate.

**Acknowledgment.** We want to acknowledge The National Polytechnic Institute (IPN, Mexico), also Center for Computing Research and especially SIP project numbers 20141310 and 20141454 for their support.

## References

1. Noll, A.M.: Shortime Spectrum and Cepstrum Techniques for Vocal Pitch Detection. Journal of Acoustical Society of America 36, 296–302 (1964)
2. John, M.: Linear Prediction: A Tutorial Review. Proceedings of the IEEE 63(4), 561–580 (1975)
3. Davis, S.B., Mermelstein, P.: Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentence. IEEE Transactions on Acoustics, Speech and Signal Processing ASSP-28(4) (August 1980)
4. Hermansky: Perceptual Linear Predictive (PLP) analysis of speech. Journal of Acoustical Society of America, 1738–1752 (April 1990)
5. Dallos, P., Fay, R.R.: Mechanics of the cochlea: modeling effects. In: de Boer, S.E. (ed.) The Cochlea. Springer, USA (1996)
6. Luis, R., Ruggero, M.A.: Mechanics of the Mammalian Cochlea'. Physiological Reviews 81(3) (July 2001)
7. Kim, D.S., Lee, S.Y., Kill, R.M.: Auditory processing of speech signals for robust speech recognition in real word noisy environments. IEEE Trans. Speech Audio Processing 7(1), 55–69 (1999)

8. Geisler, C.D.: A model of the effect of outer hair cell motility on cochlear vibration. Hear. Res. 24, 125–131 (1996)
9. Geisler, C.D., Shan, X.: A model for cochlear vibration based on feedback from motile outer hair cells. In: Dallos, P., Geilser, C.D., Matthews, J.W., Ruggero, M.A., Steele, C.R. (eds.) The Mechanics and Biophysics of Hearing, pp. 86–95. Springer, New York (1990)
10. Holmberg, M., Gelbart, D., Hemmert, W.: Automatic speech recognition with an adaptation model motivated by auditory processing. IEEE Trans. Audio, Speech, Language Processing 14(1), 44–49 (2006)
11. Haque, S., Togneri, R.: A feature extraction method for automatic speech recognition based on the cochlear nucleus. In: 11th Annual Conference of the International Speech Communication Association, InterSpeech 2010, Makuhari, Chiba, Japan, September 26-30 (2010)
12. Harczos, T., Szepannek, G., Klefenz, F.: Towards Automatic Speech Recognition based on Cochlear Traveling Wave Delay Trajectories. In: Dau, T., Buchholz, J.M., Harte, J.M., Christiansen, T.U. (eds.) 1st International Symposium on Auditory and Audiological Research (ISAAR 2007), Auditory Signal Processing in Hearing-impaired Listeners (2007) ISBN: 87-990013-1-4. Print: Centertryk A/S
13. Keener, Sneyd, J.: Mathematical Physiology. Springer, USA (2008)
14. Elliot, S.J., Ku, E.M., Lineton, B.A.: A state space model for cochlear mechanics. Journal of Acoustical Society of America 122, 2759–2771 (2007)
15. Elliott, S.J., Lineton, B., Ni, G.: Fluid coupling in a discrete model of cochlear mechanics. Journal of Acoustical Society of America 130, 1441–1451 (2011)
16. Ku, E.M., Elliot, S.J., Lineton, B.A.: Statistics of instabilities in a state space model of the human cochlea. Journal of Acoustical Society of America 124, 1068–1079 (2008)
17. Neely, S.T.: A model for active elements in cochlear biomechanics. Journal of Acoustical Society of America 79, 1472–1480 (1986)
18. Békésy: Concerning the pleasures of observing, and the mechanics of the inner ear. Nobel Lecture (December 11, 1961)
19. Mario, J.H., Rodríguez, J.L.O., Guerra, S.S., Fernández, R.B.: Computational Model of the Cochlea using Resonance Analysis. Journal Revista Mexicana Ingeniería Biomédica 33(2), 77–86 (2012)
20. Mario, J.H.: Modelo mecánico acústico del oído interno en reconocimiento de voz, Ph. D. Thesis, Center for Computing Research-IPN (Junio 2013)