

A Comparison between the Macro and Micro Mechanical Model using Values of the Human Cochlea vs. Fluid Mechanical Model for Automatic Speech Recognition

José Luis Oropeza Rodríguez and Sergio Suárez Guerra

Instituto Politécnico Nacional, Center for Computing Research,
Mexico City, Mexico

joropeza@cic.ipn.mx, ssuarez@cic.ipn.mx

Abstract. This paper shows a comparison between the macro and micro mechanical model, proposed by Neely and Kim, and extended by Elliot and Ku vs. mechanical fluid model proposed by Lesser and Berkeley both used in ASR tasks. These models are 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. Also we show an equation that show the relation distance vs. frequency extracted from the solution of the mechanical fluid model mentioned above. Neely's 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. The performance obtained was of the 98 to 100% using this model while a 98.5% performance was reached using the second propose; for a task that uses isolated digits pronounced by 5 different speakers in the Spanish language. Finally corpus SUSAS with neutral sound records with some advantages in comparison with MFCC was used.

Keywords: Speech recognition, cochlea, place theory, 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 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.

Helmholtz based his description of the function of the cochlea largely on the observations of Hensen, who had shown that the width of the basilar membrane varied along the length of the cochlea. Combining this result with Ohm’s description of sound as a combination of tones of different frequencies, led Helmholtz to the view that the cochlea performed a sort of Fourier analysis of the incoming sound waves. The physical process by which this Fourier analysis operated was the resonance of BM fibers (across

the membrane width) with different resonance frequencies (varying with the lengths of the fibers). A sound wave would set all the fibers of the BM in motion, but the ones which had resonance frequencies closest to the frequencies present in the sound signal would respond maximally. This would create a pattern over the length of the cochlea describing the frequency content of the signal: the frequency map, or place coding.

The frequency-to-place map is often characterized as a Place Theory. The specific interpretation by Helmholtz is one of the subset of Resonance Theories, which contained tuned elements, membrane resonance, or tube resonance. There was also a number of Nonresonance or Wave Theories in the Place Theory category.

The opposing category—at the time—was the group of Frequency Theories either with a nonanalytic (telephone) or analytic further specification. Or schematically:

Frequency-to-Place Map Theories around 1950

- I. Resonance Theories:
 - a) tuned elements,
 - b) membrane resonance,
 - c) tube resonance.
- II. Frequency Theories:
 - a) analytic,
 - b) nonanalytic (e.g., telephone theory).

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. This expression is then compared with the macro and micro mechanical model proposed by Nelly&Kim in 1986, with the objective to analyze the performance in a Automatic Speech Recognition task for two databases. First of them we use Spanish digits and the second using neutral SUSAS Corpus.

In the micromechanical the anatomical structure of a radial cross-section (RCS) of the cochlear partition (CP) is illustrated in the following figure 1. 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.

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]:

$$\frac{d^2}{dx^2} P_d(x) = \frac{2\rho}{H} \ddot{\varepsilon}(x) \quad (1)$$

$$\frac{d^2}{dx^2} P_d(0) = 2\rho \ddot{\varepsilon}_s \quad (2)$$

$$\frac{d^2}{dx^2} P_d(0) = 2\rho \ddot{\varepsilon}_s \quad (3)$$

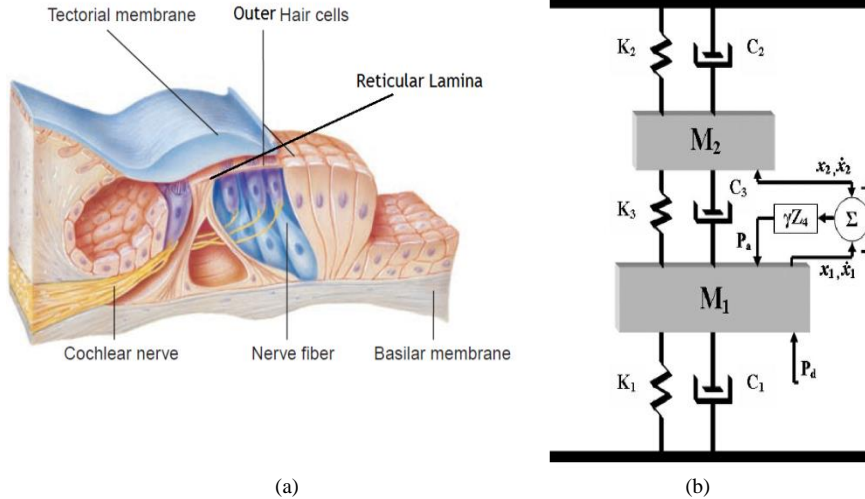


Fig. 1. Anatomical structure of the cochlear partition (a). The outer hair cells, micro mechanical representation (b).

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.

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

For another side, we have a model proposed by Lesser and Berkley in 1972 that employs mechanical fluids to solve the same problem. Let $u = (u_1, u_2, u_3)$ be the fluid velocity, p the pressure, and ρ the constant density of the fluid. The mass of fluid in a fixed volume V can change only in response to fluid flux across the boundary of the volume. Thus [23],

$$\frac{d}{dt} \int_V \rho dV = - \int_S \rho (u \cdot n) dS = 0 \tag{4}$$

where S is the surface of V , and $n = (n_1, n_2, n_3)$ is the outward unit normal to V .

After considering that the momentum of the fluid in a fixed domain V can change only in response to applied forces or to the momentum flux across the domain boundary, and using the divergence theorem to convert surface integrals to volume integrals, 2 is obtained:

$$\int_V \left(\rho \frac{\partial u_i}{\partial t} + \rho \nabla \cdot (u_i u) + \frac{\partial p}{\partial x_i} \right) dV = 0 \tag{5}$$

After considering that V is arbitrary, fluid motions are of small amplitude and there is an irrotational flow, the following equations are shown:

$$\begin{aligned} \rho \frac{\partial \phi}{\partial t} + p &= 0, \\ \nabla^2 \phi &= 0 \end{aligned} \tag{6}$$

Lesser and Berkley developed a model that combines these last two equations with the equation of a damped, forced harmonic oscillator and is considered one of the simplest of the cochlea models. They propose that each point of the basilar membrane is modeled as a simple damped harmonic oscillator with mass, damping, and stiffness that vary along the length of the membrane.

Thus, the movement of any part of the membrane is assumed to be independent of the movement of neighboring parts of the membrane, as there is no direct lateral coupling. The deflection of the basilar membrane, $\eta(x, t)$, is specified by a model of a forced harmonic oscillator defined as

$$m(x) \frac{\partial^2 \eta}{\partial t^2} + r(x) \frac{\partial \eta}{\partial t} + k(x)\eta = p_2(x, \eta(x, t), t) - p_1(x, \eta(x, t), t) \tag{7}$$

where $m(x) = 0.1$, $r(x) = 300 e^{-ax}$, $k(x) = 10^9 e^{-2ax}$. An analytical solution of this problem can be found using standard Fourier series [23]. Solutions of this form are looked for:

$$\phi = x \left(1 - \frac{x}{2} \right) - \sigma y \left(1 - \frac{y}{2\sigma} \right) + \sum_{n=0}^{\infty} A_n \cosh[n \pi (\sigma - y)] \cos (n \pi x) \tag{8}$$

This paper proposes solving the Lesser and Berkley equation using the solution proposed in [20]. This solution is related with the place theory of hearing, initially proposed by Von Békésy. To perform the analysis each section of the membrane is considered as a forced harmonic isolated oscillator, which is excited by an external force $F e^{j\omega t}$ that represents the driving force on each section of the basilar membrane and this force is produced by vibrations transmitted into the cochlea by the oval window. Two solutions are proposed related with the before mentioned equation. Firstly, the forced harmonic oscillator is represented by the following equation

$$m(x) \frac{d^2 \eta}{dt^2} + R_m(x) \frac{d\eta}{dt} + k(x)\eta = F e^{j\omega t} \tag{9}$$

where m is the mass, R_m mechanical resistance and k is the damping constant.

Considering that $\eta = A e^{j\omega t}$, then amplitude of the wave sound into the cochlea is represented by [20]. Secondly, a damped harmonic oscillator with the following equation is considered:

$$m(x) \frac{d^2 \eta}{dt^2} + R_m(x) \frac{d\eta}{dt} + k(x)\eta = 0 \tag{10}$$

Then, a solution is given by

$$\eta = A e^{-\beta t} \cos (\omega_0 t + \phi) \tag{11}$$

Equation 12 shows that the amplitude for each section of the membrane depends of the frequency ω in the applied force. The amplitude has a maximum when the denominator has its minimum value and this occurs at a specific frequency excitation

called resonance frequency. This is defined by the values of mass and stiffness, when the frequency ω of the applied force is equal to $k(x)/m(x)$ it is said that the system is resonant in amplitude and obtains the maximum value of the basilar membrane displacement. This last equation can be expressed as a function of frequency and distance, if considering that $\omega = 2\pi f$ thus, this is possible using our purpose literature does not find an equal relationship [20]:

$$A = \frac{F / m(x)}{\sqrt{\left(4\pi^2 f^2 - \frac{k(x)}{m(x)}\right)^2 + 4\pi^2 f^2 \frac{R_m(x)^2}{m(x)^2}} \quad (12)$$

Figures 2 and 3 show the behavior of the basilar membrane with the values obtained when we calculate the equation obtained. 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. We don't use an analytical expression to obtain the response of Neely model.

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.

$$d(n) = d_{\max} + \sum_{n=0}^{n_{\text{int}}} n \frac{d_{\min} - d_{\max}}{n_{\text{int}} + 1} \quad (13)$$

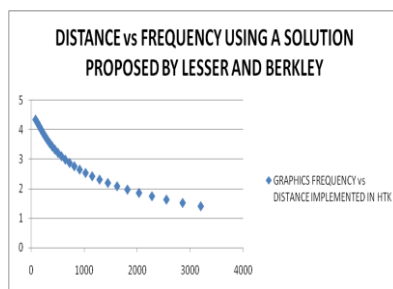


Fig 2. Distance vs Frequency using model Lesser & Berkley's

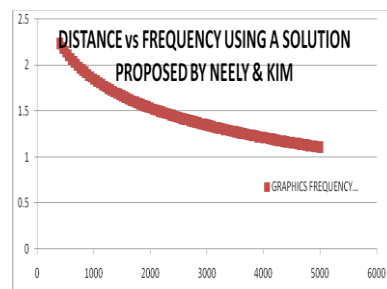


Fig 3. Distance vs Frequency using model Neely's

In [5] 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 the proposal in [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.
 - 2.1 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
 - 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
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 1. 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
EPCC RESONANCE ANALYSIS	99.25	99.35	99.6	EPCC RESONANCE ANALYSIS	99.35	99.45	99.75

Table 2. LPC, CLPC, MFCC, DELTA and ACCELERATION coefficients

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
EPCC RESONANCE ANALYSIS	99.30	99.6	99.7	EPCC RESONANCE ANALYSIS	99.45	99.75	99.8

Table 3. 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
EPCC RESONANCE ANALYSIS	99.4	99.6	99.8	EPCC RESONANCE ANALYSIS	99.6	99.8	99.8

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 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 4,410 files of speech were processed. Finally, Tables 4 shows 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.

Table 4. Results obtained using HTK, SUSAS Corpus and manual labeling

	MFCC		EPCC Using Neely values		EPCC Using Ku values		EPCC Using Elliot values		EPCC Using resonance analysis	
	sent.	word	sent.	word	sent.	word	sent.	word	sent.	word
<i>boston1</i>	91.84	92.06	90.61	90.87	90.2	90.48	89.39	89.68	90.2	90.84
<i>boston2</i>	95.51	95.63	93.47	93.65	93.47	93.65	93.06	93.25	93.88	94.05
<i>boston3</i>	96.73	96.83	93.88	94.05	95.92	96.03	96.33	96.43	92.65	92.86
<i>general1</i>	96.73	96.83	92.24	92.46	93.88	94.05	93.88	94.05	95.51	95.24
<i>general2</i>	94.29	94.44	90.61	90.87	90.61	90.87	89.39	89.68	93.06	93.25
<i>general3</i>	93.47	93.65	88.16	88.49	93.47	93.65	93.06	93.25	94.69	94.84
<i>nyc1</i>	91.84	92.06	91.84	91.67	87.35	87.3	96.33	96.43	93.06	92.86
<i>nyc2</i>	91.02	91.27	91.84	92.06	86.53	86.9	93.88	94.05	89.8	90.08
<i>nyc3</i>	95.92	96.03	92.65	92.86	90.61	90.87	89.39	89.68	90.2	90.48

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. The interest of this paper is show the implementation of the cochlear models in Automatic Speech Recognition tasks. We show that the theory of these models can be used to obtain parameters from the speech signal and used as input to the Hidden Markov Model Toolkit.

Also, the paper show an analytic solution to the Lesser & Berkley model (this model was proposed in 1972 and is based in the mechanical fluid and its solution used the Fourier series), that is based in the resonance analysis proposed by Helmholtz. Then we show a mathematical expression can be compared with another used in the State of the Art, for example the equation of Greenwood. This article demonstrated that our propose is very interesting because the performance reached was adequate and 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), Center for Computing Research and especially SIP project 2015104 for their support.

References

1. Noll A. M.: Shorttime Spectrum and Cepstrum Techniques for Vocal Pitch Detection. *Journal of Acoustical Society of America*, Vol. 36, pp. 296–302 (1964)
2. Makhoul John: Linear Prediction: A Tutorial Review. *Proceedings of the IEEE*, Vol. 63, No. 4, pp. 561–580 (April 1975)
3. Steven B. Davis, Paul Mermelstein: Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentence. *IEEE Transactions on Acoustics, speech and signal processing*. Vol. ASSP-28, No. 4 (August 1980)
4. Hermansky H.: Perceptual Linear Predictive (PLP) analysis of speech. *Journal of Acoustical Society of America*, pp. 1738–1752 (April 1990)
5. de Boer S.E.: *Mechanics of the cochlea: modeling effects*. The cochlea. Springer (USA) (1996)
6. Robles Luis, Ruggero Mario A.: *Mechanics of the Mammalian Cochlea*. *Physiological Reviews*, Vol. 81, No. 3 (July 2001)
7. D. S. Kim, S. Y. Lee, R. M. Kill: Auditory processing of speech signals for robust speech recognition in real word noisy environments. *IEEE Trans. Speech Audio Processing*, vol. 7, no. 1, pp. 55–69 (Jan 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. *The Mechanics and Biophysics of Hearing*, edited by P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, C. R. Steele, Springer-Verlag, New York, pp. 86–95 (1990)
10. M. Holmberg, D. Gelbart, W. Hemmert: Automatic speech recognition with an adaptation model motivated by auditory processing. *IEEE Trans, Audio, Speech, Language Processing*, vol. 14, no 1, pp. 44–49 (Jan. 2006)
11. S. Haque, R. Togneri: A feature extraction method for automatic speech recognition based on the cochlear nucleus. *InterSpeech 2010, 11th Annual Conference of the International Speech Communication Association*, Makuhari, Chiba, Japan, September 26-30 (2010)
12. T. Harczos, G. Szepannek, F. Klefenz: Towards Automatic Speech Recognition based on Cochlear Traveling Wave Delay Trajectories. *Auditory signal processing in hearing-impaired listeners, 1st International Symposium on Auditory and Audiological Research (ISAAR 2007)*, T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.) (2007)
13. Keener J.P., Sneyd J.: *Mathematical Physiology*. USA: Springer (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*, Vol. 130, pp. 1441–1451 (2011)
16. Ku EM, Elliot SJ, Lineton BA.: 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*, Vol. 79, pp. 1472–1480 (1986)
18. Békésy, G.: Concerning the pleasures of observing, and the mechanics of the inner ear. *Nobel Lecture*, December 11 (1961)

19. Mario Jiménez Hernández, José Luis Oropeza Rodríguez, Sergio Suárez Guerra, Ricardo Barrón Fernández: Computational Model of the Cochlea using Resonance Analysis. *Revista Mexicana Ingeniería Biomédica*, Vol. 33, Num. 2, pp. 77–86 (diciembre 2012)
20. Mario Jiménez Hernández: Modelo mecánico acústico del oído interno en reconocimiento de voz. Ph. D. Thesis, Center for Computing Research-IPN (junio 2013)