Biol. Cybern. 77, 123–130 (1997)

Biological Cybernetics

Artificial implementation of auditory neurons: A comparison of biologically motivated models and a new transfer function oriented model

U. Meyer-Bäse[1,2], **H. Scheich**[2]

[1] Institute of Digital Technics, TU Darmstadt, D-64283 Darmstadt, Germany
[2] Federal Institute for Neurobiology, D-39008 Magdeburg, Germany

Received: 26 March 1997 / Accepted in revised version: 27 May 1997

Abstract. Auditory perception neurons, also called inner hair cells (IHCs) because of their physical shape, transform the mechanical movements of the basilar membrane into electrical impulses. The impulse coding of the IHC is the main information carrier in the auditory process and is the basis for improvements in cochlear implants as well as for low-rate, high-quality speech processing and compression. This paper compares biologically motivated models (Meddis, Cooke, Payton) with a newly developed model which is transfer function oriented. The new model has only three reservoirs and the parameters can be controlled through five small ROM tables. This model is compared with the often used Meddis model in terms of accuracy, system parameter flexibility, and hardware effort in an FPGA implementation.

1 Introduction

First a short review of ear physiology is given to establish the central rule of inner hair cells (IHCs) in the hearing process. The hearing process in humans is a complex multistage system (Handel 1989; Zwicker and Fastl 1990, Furui and Sondhi 1993) as seen in Fig. 1 (Homes 1988). The sound impinging on the side of the head travels down the auditory channel to reach the tympanic membrane. In the middle ear the movements of the tympanic membrane are coupled with the oval window at the end of the cochlea. The cochlea is the main structure of the inner ear. The waves travel along the basilar membrane, which shows frequency-selective movements (Békésy 1942). The auditory neurons which, because of their physical shape, are also called hair cells, convert the movements to neural responses. There are about 30 000 inner and outer hair cells; the outer hair cells can sharpen the frequency selectivity. By means of very small movements of the basilar membrane, neural responses in the brain are produced. This neural response shows both intensity coding and phase locking. The parameters of the intensity coding are short-term firing rate, long-term firing rate, and the refraction time (minimum time between two pulses). Figure 2 shows the response of an IHC to a sudden sinus impulse. It can be seen that the short-term rate is much higher than the long-term rate.

Now consider the temporal pattern of the nerve impulses. Recall that points on the basilar membrane will oscillate at the frequency of the driving vibration. At one extreme, the nerve fiber can fire randomly throughout the basilar membrane oscillation, and at the other extreme the nerve fiber can fire at only one point in the movement. In this latter case, the firings are *phase-locked* to one point of the sound wave. One way to display the firing pattern is to measure the interval between a fixed point on the stimulating wave and the first spike modulo, the period length T of the stimulating wave. The number of spikes occurring in each interval are cumulated to create the *period histogram.* The abscissa of the period histogram is the time between the fixed point of the stimulating wave (typically, upward zero-crossing) and the first spike modulo T; the ordinate shows the number of spikes in each such interval. If spikes tend to occur at one point of the waveform, then the nerve fibers are called *in-lock* (e.g., Fig. 3a, 93 dB); otherwise, for a uniform distribution, they are *out-of-lock* (e.g., Fig. 3a, 0 dB).

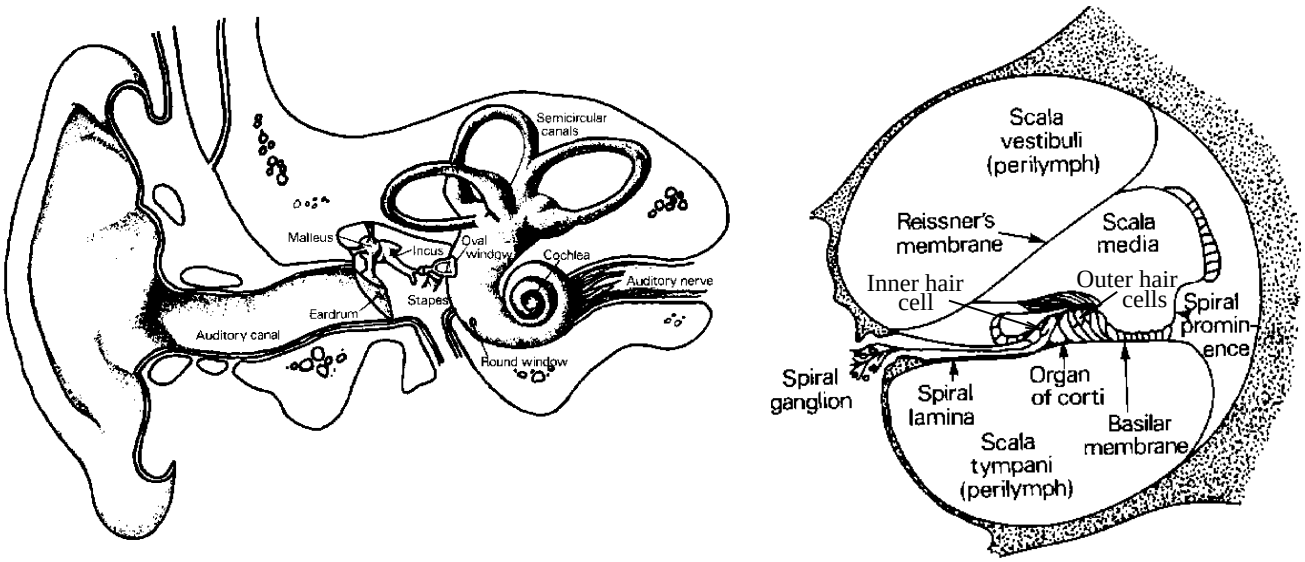
The phase locking can be judged with the synchronization index after Anderson (1973):

$$\text{Index} = 1/N \times \sqrt{\left(\sum_{x=0}^{T-1} h[x]\sin(2\pi x/T)\right)^2 + \left(\sum_{x=0}^{T-1} h[x]\cos(2\pi x/T)\right)^2} \quad (1)$$

$$\text{where } N = \sum_{x=0}^{T-1} h[x],$$

where T is the period of the stimulating signal. It can be seen from (1) that the index computes the first harmonic in the period histogram. Figure 4 shows the period histogram (bottom) for a period length of 1 ms, and the two computed values for the Anderson index by stimulating with a sum of 1 and 4 kHz signal for eight IHCs with different characteristic frequencies. Note that for two stimulating waves, also two period histograms (T_1 = 1 ms, T_2 = 0.25 ms) must be generated to compute the correct index with (1).

Correspondence to: U. Meyer-Bäse (Tel: +49-6151-163775, e-mail: umb@marlowe.dtro.e-technik.th-darmstadt.de)



(a) (b)

Fig. 1. a Structure of the human auditory system. **b** Cross-section of one turn of the cochlear spiral

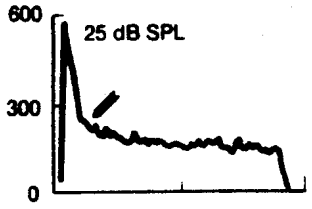


Fig. 2. Firing rates of an inner hair cell following a sudden sinus impulse (Handel 1989)

A second way to display the temporal firing pattern is to measure the time between each pair of successive spikes without regard to the stimulating wave. These *interspike* intervals create an *interspike interval histogram (ISI)* in which the abscissa represents the time interval between successive spikes and the ordinate represents the number in each such intervals. The ISI histogram also provides a display of the refraction time. Figure 3b shows biological measurement of the ISI histogram, while Fig. 4 (top) shows a computer simulation for eight channels.

So far it is not totally clear whether the main information in the auditory process is transported by intensity coding and/or phase-locking from the IHCs to the cochlear nucleus. The investigation by Sachs et al. (1982) indicated that phase-locking dominates at higher sound levels. Nevertheless a good IHC model should preserve both: intensity coding and phase-locking.

2 Inner hair cell models

From an information processing standpoint the IHCs are the main processing elements in the auditory process. The impulse coding of the IHC can be used to improve cochlear implants, which have been used recently for deaf patients (Working Group for Hearig-Impaired 1991), and in low-rate, high-quality speech processing and compression (Ghitza 1994). Therefore, several research teams have tried to build such auditory neurons (Hewitt and Meddis 1991). These models differ in functionality and complexity. Table 1 shows that from the commonly known models only the Payton and Meddis models show both intensity coding and phase-locking. The complexity advantage of the Meddis model motivated its use as a reference.

Table 1. Comparison of different IHC models (Hewitt and Meddis 1991)

Model	No. of reservoirs	No. of parameters	Short- and long-term adaptation	Phase-synchro-nization
Schroeder	1	3	No	No
Oono	1	3	Yes	No
Allen	1	10	No	Yes
Schwid	6	15	Yes	No
Ross	4	9	Yes	No
Payton	1024	19	Yes	Yes
Cooke	3	6	Yes	No
Meddis	3	8	Yes	Yes
Meyer–Bäse	3	(5 ROMs)	Yes	Yes

2.1 The Meddis model

First, the Meddis model was realized in the object oriented simulation framework Ptolemy provided by the Berkeley group of Prof. Messerschmitt through the WWW. Figure 4 shows ISI histograms, and the period histogram for computing the Anderson index for a computer simulation of the Meddis model with an input signal consisting of the sum of 1 and 4 kHz signals. A Stanford-Implant eight-filter model having logarithmic coverage of the frequency range from 900 to 8000 Hz was used. Figure 5 shows the simulation results for different levels of the input signal for two different parameter sets. Note that the synchronization index has a higher sensitivity (lower threshold) and reaches much earlier saturation than the ‘onset’ and ‘steady-state’ rate functions.