Coding of Auditory Information into Nerve-Action Potentials

W. Hemmert, M. Holmberg, M. Gerber
Infineon Technologies AG Munich, Corporate Research, Email: werner.hemmert@infineon.com

Introduction

Neuronal sound processing requires compression of the huge dynamic range of acoustical signals to the limited range which can be coded by neurons. Here we present a compressive inner ear model followed by a realistic inner hair cell model. Spike trains transmitted by the auditory nerve fibers allow for spectral, temporal and spatial information processing, which is required for robust speech recognition, especially in noisy environments.

Modeling

Sound signals were delivered to a simplified model of the eardrum which transferred sound pressure into motions of the middle ear, which were fed into the inner ear model.

Inner ear model with large dynamic compression

Basilar membrane (BM) vibrations were calculated with a computationally efficient wave-digital filter model, consisting of a bank of mass-spring resonators (125) connected by a coupling-mass[1,2]. The frequency map of the resonators was adjusted according to Greenwood's[3] map for the human inner ear. The wave-digital filter model solves the passive inner ear hydrodynamics (one-dimensional case), which describes the vibration of the BM at high levels (compare Fig. 1). Inner-ear hydrodynamics is responsible for the frequency-place transformation of sound in the inner ear: High frequency sounds cause maximal vibration amplitudes at the base of the cochlea, whereas low-frequency stimuli travel further towards the cochlear apex. In the living hearing organ, low-level sounds are mechanically amplified, probably by the so-called outer hair cells (OHCs), which are thought to sense the hearing organ's motion and feed back mechanical energy into its vibration. This active amplification stage both boosts the vibration amplitudes at low sound levels and significantly sharpens the traveling-wave at low levels. Recent measurements have shown that the amplification may well be more than thousand-fold (>60 dB, [4]). As the nature of the amplifier is still unknown and as the implementation of a feed-back amplifier with such a high gain poses severe stability issues, we have chosen a different implementation. Instead of amplifying the BM response, we added second-order resonators at the outputs of the cochlear filter bank and modulated their quality factors. Note that these passive resonators amplify only the vibration amplitude, not the energy of the vibration, as it is the case in the living cochlea. Quality factors were altered in every iteration-step depending on the instantaneous displacement of each resonator, using a Boltzmann-function similar to the sensitivity function of the OHCs. By cascading multiple resonator stages (four) and modulating their quality factors from ten to one, we achieved large amplification (up to 80 dB) together with reasonable filter shapes.

Fig. 1 shows the "amplified" BM response along the cochlea for a 1 kHz tone presented at various input levels. At high input levels, responses share the characteristics of a passive traveling-wave with a gradual build-up starting from the cochlear base and a sharp roll-off after the relatively broad maximum is reached (Fig. 1, green line). At low levels, filter shapes are narrow and almost symmetrical. The amplitudes at the most sensitive location (at the characteristic frequency, CF) are greatly amplified, so that even faint sounds cause excitation above threshold (rate-threshold was estimated between 1.5 nm and 100 µm/s [4]). At the most sensitive place, the growth function of BM displacement is compressed and follows approximately a cube root law at medium levels. At more basal locations growth functions are almost linear. For increasing intensities, the excitation pattern grows highly asymmetrical, which is well-known as the upper spread of masking.

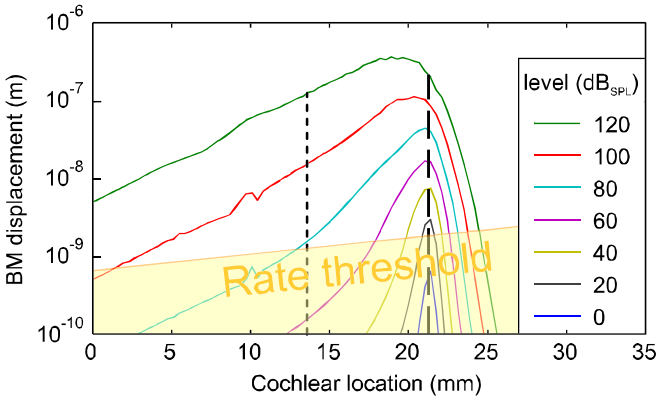

Fig. 1: Model-excitation pattern (RMS) for a 1 kHz pure tone. Note that responses are greatly compressed at the characteristic location (dashed line 21 mm) and almost linear at more basal locations (dotted line). The yellow line indicates rate threshold of the auditory nerve.

Inner hair cell model

Basilar membrane vibrations drive the stereociliary bundles of the sensory cells, the inner hair cells (IHC, Fig. 2) by fluid motion. In the model, fluid velocity (v_{fluid}) around the bundles was assumed to be equal to BM motion. Fluid friction and hair bundle stiffness form a high-pass filter with a corner frequency in the range of 1 kHz, which sharpens the low-frequency input to the IHC. When the hair bundle is deflected, ion channels open at its tip and a transduction current, mainly carried by K^+-ions, is driven by difference between endocochlear potential (EP, +80 mV) and membrane potential (U_M, -40 mV) into the cell. The mechano-electrical transduction can be described by a double-Boltzmann function [5]. This function shows saturation for displacements larger than about 100 nm. The saturation severely limits the dynamic range of IHCs to a value around 40 dB. The transduction current depolarizes the IHC membrane, which was modeled by its capacitance (10 pF) and the bulk conductivity of its ion channels (60 nS). Upon depolarization of the membrane, volt-

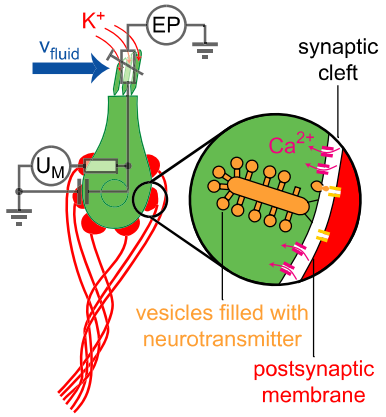

Fig. 2: Schematic view of the inner hair cell with its synaptic region magnified.

age dependent Ca-channels close to the cell's synaptic terminals are activated and Ca^{2+}-ions enter the cell. Elevated Ca^{2+}-levels cause fusion of synaptic vesicles with the cell membrane. The neurotransmitter contained in the vesicles diffuses across the synaptic cleft, binds to the receptors in the postsynaptic membrane which is depolarized and a nerve action potential is propagated along the auditory nerve fiber (ANF) to the brain. Vesicle fusion is dominated by a so-called readily releasable pool (RRP) of vesicles, which are located in the synaptic region closely to the cell membrane (see Fig. 2). A large stimulus causes the fusion of many vesicles. Due to the depletion of the RRP, in the following few tens of milliseconds the spiking probability of the auditory nerve is reduced, an effect which is known as adaptation. The kinetics of vesicle fusion and RRP refill was modeled according to recent measurements[6,7]. In addition to adaptation, an ANF can not fire twice within an interval of approximately 1 ms. This refractory process was modeled according to data from Carney[8]. As the process of vesicle fusion is a statistical process, also the generation of nerve-action potentials is random. Fig. 3 shows the excitation pattern of the auditory fibers along the cochlea to an artificial vowel. Only one ANF per section was modeled and plotted. Responses are delayed towards the apex due to propagation of the traveling wave and due to the additional nonlinear resonators. The two formants of the artificial vowel, not individual harmonics, are separated along the cochlea. The fundamental frequency of the vowel (100 Hz corresponding to 10 ms) is clearly visible as an amplitude modulation, especially at the second formant.

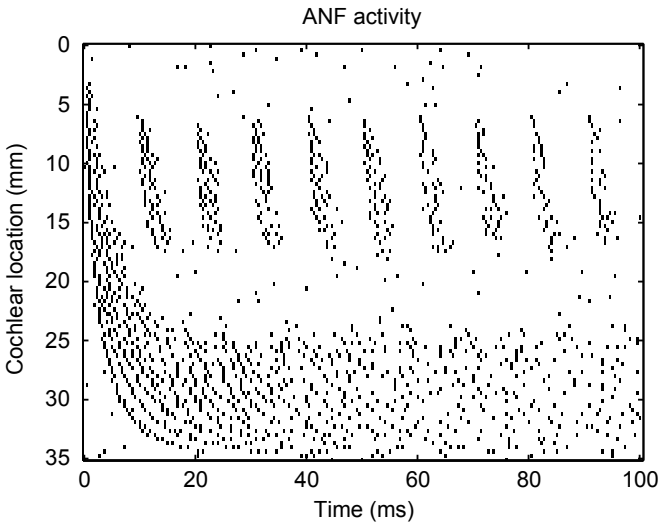

Fig. 3: Firing pattern of auditory nerve fibers along the cochlea to an artificial vowel "ε". Note that both spectral and temporal features of the sounds are conserved. Especially the fundamental frequency of 100 Hz is apparent as an amplitude modulation at the location of the second formant.

Efferent feedback on inner ear mechanics

The nonlinear compression of the inner ear is also under efferent control via the superior olivary complex (SOC, Fig. 4). Efferent nerve fibers contact the OHCs and a reduction of the sensitivity of the auditory nerve by 20 dB was observed when they are activated[9]. For computational reasons, not actual action potentials but spiking probabilities of the ANF was integrated over an area of about an octave, low-pass filtered (τ=200 ms) and allowed to reduce the amplification of the BM by maximal 20 dB at the highest neuronal activity. The action of the efferent feed-back for a 5 kHz CF tone is illustrated in Fig. 5. With the efferent system active, the growth function at the characteristic place is depressed with growing intensities. If the same location is stimulated by a 3.5 kHz tone, well below CF, the growth function of the auditory nerve's firing rate has a much higher threshold and saturates already within about 20 dB, not effected by the efferent system.

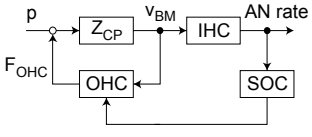

Fig. 4: Schematics of the influence of the medial efferent system on inner ear amplification.

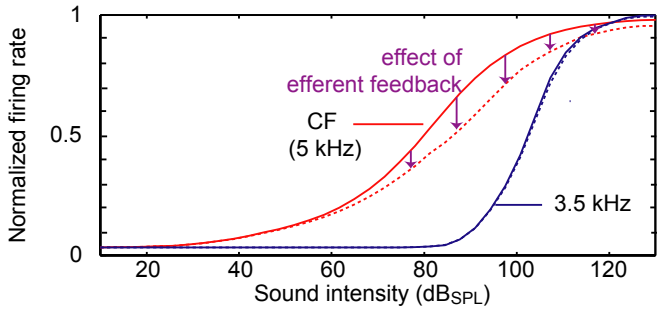

Fig. 5: Auditory nerve growth-function (5 kHz characteristic location) for pure tone stimuli at CF (red lines) and below CF (blue lines). With feedback from the medial efferent system (dotted lines) the activity at CF is depressed.

Conclusions

We have developed a phenomenological model of the inner ear which achieves large dynamic compression of BM vibration as observed in the hearing organs of mammals. This compression is required to code sound signals into nerve-action potentials. A neuronal feed-back mechanism corresponding to the lateral efferent system was modeled which finely controls BM-vibration and improves the signal-to-noise ratio in the presence of background noise. Due to its details, our model is suitable to examine the effects of damage to OHC, IHC and the ANF. At Infineon Technologies, we are interested to study coding of acoustical information into nerve action potentials and to extract features in both spectral- and temporal domains, which is essential as temporal information is required for robust speech recognition in noisy and reverberant environments.

[1] Strube HW: A computionally efficient basilar-membrane model. Acustica 1985 (58), p. 207-214.

[2] Zwicker E: A hardware cochlear nonlinear preprocessing model with active feedback. JASA 1986: 146-153.

[3] Greenwood DD: A cochlear frequency-position function for several species-29 years later. JASA 1990: 2592-2605.

[4] Ruggero MA, Narayan SS, Temchin AN, Recio A: Mechanical bases of frequency tuning and neural excitation at the base of the cochlea: comparison of basilar-membrane vibrations and auditory-nerve-fiber responses in chinchilla. PNAS 2000: 11744-11750.

[5] Mountain DC, Cody AR: Multiple modes of inner hair cell stimulation. Hear.Res.1999: 1-14.

[6] Beutner D, Voets T, Neher E, Moser T: Calcium dependence of exocytosis and endocytosis at the cochlear inner hair cell afferent synapse. Neuron 2001: 681-90.

[7] Moser T, Beutner D: Kinetics of exocytosis and endocytosis at the cochlear inner hair cell afferent synapse of the mouse. PNAS 2000: 883-888.

[8] Carney LH: A model for the responses of low-frequency auditory-nerve fibers in cat. JASA 1993: 401-417.

[9] Guinan JJ, Stankovic KM: Medial efferent inhibition produces the largest equivalent attenuations at moderate to high sound levels in cat auditory-nerve fibers. JASA 1996: 1680-90.