How Onset Neurons Process Speech Sounds

Werner Hemmert, Marcus Holmberg, Christoph K. Schmid, Ulrich Ramacher
Infineon Technologies AG, Corporate Research, 81730 München, Germany,
Email: {marcus.holmberg,werner.hemmert,ulrich.ramacher}@infineon.com

Introduction

Octopus neurons are located in the cochlear nucleus, the first neuronal processing stage. They receive direct inputs from auditory nerve fibers (ANF) and code temporal information with high precision, which is essential for sound localization and probably also for the perception of pitch. Unlike most other neurons, which exhibit sustained activity to continuous excitation, octopus neurons show onset inhibitory responses: they only fire on signal onsets. As they receive predominantly excitatory inputs, their membrane properties are thought to be responsible for suppressing sustained responses [1],[3],[4]. Octopus neurons are also are sensitive to amplitude modulations. Their processing strategy is interesting because they suppress spontaneous- and uniform activity, a feature which is essential for sound processing in noise.

Modeling

In this paper we combine our realistic inner ear model [2] with a model of so-called onset inhibitory units, which have been identified as octopus neurons. They receive excitatory inputs from approximately 60 auditory nerve fibers and they fire if about 10% to 25% of their inputs are activated within 1 ms [1]. In this study, we only used output trains from high spontaneous rate (HSR) ANFs.

We established a single-compartment octopus model with Hodgkin-Huxley-type ion channels. Rothman and Manis measured the properties of their five major conductances [3] and derived both steady-state and dynamic equations. We solved the differential equations by iteration in the time domain and using conductance values and time constants corrected for a body temperature of 38°C. Octopus neurons – and our model – exhibit large activated ionic conductances at rest of about 40 nS and a membrane capacitance of 12 pF. Due to their extraordinarily fast membrane time constant (0.3 ms) they act as coincidence detectors and they greatly enhance the precision of timing compared to a single ANF. The electrical behavior of octopus neurons is dominated by a low-threshold potassium channel (K_{LT}) with activation kinetics in the order of 2 ms, which is already activated at rest [4]. When their membrane is depolarized, they elicit an initial action potential, but thereafter K_{LT} activates and counteracts input currents which keeps the membrane potential below spiking threshold.

In this study we connected 60 auditory nerve fibers originating from a single frequency channel (characteristic frequency: 1.4 kHz) of our model with an octopus neuron. We used only excitatory synapses. Activation by an action potential of the auditory nerve was modeled with a single exponentially decaying conductance (0.1 ms decay-time constant, [3]). We set the peak value to 8.5 nS; an action potential was elicited when 25% of the input fibers fired synchronously.

Results

The firing rate of ANFs cover only a very small range; sustained rates of HSR fibers are between 30 and 300 spikes/s. Sound signals, on the other hand, can cover a huge dynamic range of more than 6 orders of magnitude! To overcome this discrepancy, dynamic compression is a key function performed by our inner ear. Our model provides large dynamic compression of more than 60 dB, which enables neuronal processing over a wide dynamic range. Fig. 1 shows HSR fibers coding a 1.5 kHz pure tone with stepwise increasing amplitude; sound pressure levels were 40, 60 and 80 dB (rms). The

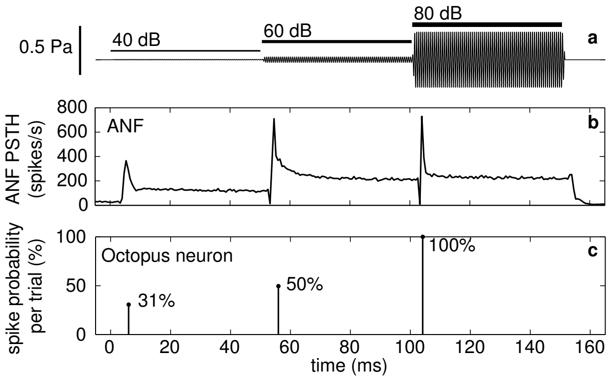

Figure 1: Onset processing of octopus neurons. (a) sound stimulus is a 1.5 kHz pure tone with stepwise increasing amplitude (rise time: 1 ms). (b) Poststimulus-time histogram (PSTH) of a single ANF (0.66 ms time bins). (c) octopus neuron activity per trial (100 repetitions).

top trace shows the signal. Note that signal amplitude at 40 dB (starting at $t = 0$ s) is a factor of 100 smaller than at 80 dB. The reaction of an ANF with a characteristic frequency of 1.4 kHz, slightly lower than the test tone, is displayed in the middle trace. At rest ($t < 0$ s) the ANF fires with its spontaneous rate of approximately 30 spikes/s. Upon signal onset, the ANF reacts with a sharp rise of its firing rate which decays to a sustained rate of approximately 120 spikes/s. For a tenfold increase in stimulus amplitude (60 dB, $t = 50$ ms) another transient is generated and the spontaneous rate increases to 210 spikes/s. For yet another tenfold amplitude increment the sustained firing rate hardly increases any more

(220 spikes/s); HSR fibers saturate about 40 dB above threshold. Still, the fiber is able to generate a transient onset response.

Temporal processing by octopus neurons

For high-frequency sounds, octopus neurons fire only at signal onsets, whereas for frequencies up to about 800 Hz, they can fire at every cycle. Fig. 1c shows the reaction of an octopus neuron to a tone with stepwise increasing amplitude. At the onset of the 40 dB tone, the neuron fires with 31% probability (100 trials). During the duration of the tone-burst, K_{LT} activates and prevents further action potentials. This mechanism still works when the signal amplitude is increased 10-fold (even twice). When the signal level increased from 40 to 60 dB, the neuron fired with a probability of 50%, for the increase to 80 dB an action potential was always elicited. Thereby the temporal precision of the action potentials is remarkable: while ANFs fire for the whole half stimulus cycle, 90% of the action potentials of the octopus neuron were in an interval of less than 150 μs. This precision is reached by coincidence detection of at least four synchronously firing ANFs and by the fast membrane time constant of octopus neurons. The extraction of signal onsets by octopus neurons probably also plays an important role in speech perception. Octopus neurons are preferentially triggered by stop-consonants (data not shown) and by voiced speech.

The spike-triggered reverse-correlation technique revealed the temporal excitation pattern which most likely causes the octopus neuron to fire [4]. The reverse correlation can be interpreted in a similar way as the impulse response of a linear filter. Its spectrum revealed that temporal processing of octopus neurons very much resembles a band-pass filter. The low-frequency slope was close to 6 dB/oct, the pass-band (-3 dB) reached from 110 Hz to about 650 Hz (data not shown). This finding indicates that octopus neurons might also play a role in the processing of amplitude modulated (AM) sounds. We tested the response of ANFs and octopus neurons to AM tones (data not shown) and to speech sounds. Vowels[1] are characterized by periodic glottis strokes (which defines their fundamental frequency). This periodicity is clearly visible in the signal's time trace (Fig. 2a). Fig. 2b displays the PSTH of ANFs with a characteristic frequency of 1.4 kHz, at the location of the second formant of the vowel [ɝ] (as in "*bird*"). Also in this frequency range the periodicity is clearly visible; ANFs preserve the temporal structure of sounds. Octopus neurons extract the periodicity with great precision. They fire very reliably at signal onset (firing probability: 100%, determined from 100 trials) and with high temporal precision (spikes were within a 250 μs window) at almost each glottis stroke (firing probability was always higher than 96%) and there is hardly any intermediate activity (Fig. 2c).

[1]We used artificial vowels generated by a Matlab implementation of the Klatt80 synthesizer, which can be found at http://www.dal.ca/~mkiefte/klatt/.

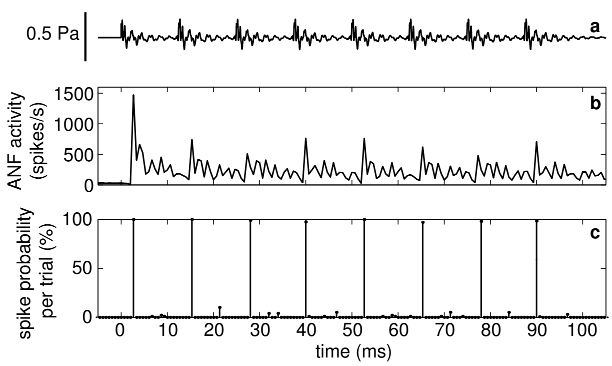

Figure 2: Processing of an artificial vowel [ɝ] as in "*bird*". (a) sound stimulus (80 Hz fundamental frequency; formant frequencies: F1=490 Hz, F2=1.4 kHz, F3=3.7 kHz), (b) PSTH of a single ANF (CF=1.4 kHz, 0.66 ms time bins) and (c) activity of an octopus neuron (CF=1.4 kHz, firing probabilities determined with 100 repetitions).

Conclusion

Auditory nerve fibers of the auditory system code both spectral- and temporal properties of sound. Temporal precision is greatly enhanced by octopus neurons, which require synchronous firing of multiple ANFs within a brief (<1 ms) period. Octopus neurons play a major role in sound localization, they enhance temporal precision and fire at signal onsets with a jitter of less than 150 μs. They suppress continuous activity over a wide amplitude range and still fire at novel signal onsets. The spectrum of the spike-triggered reverse correlation shows the band-pass behavior of octopus neurons. Octopus neurons are most sensitive in a frequency range between 110 Hz and 650 Hz. This range covers the periodicity of voiced speech, which is preserved in the temporal responses of the ANFs along the cochlea. Octopus neurons extract the periodicity of vowels with great precision, they usually fire only once for each period. It is therefore likely that these neurons not only contribute to spatial hearing, but also play an important role in speech processing.

Acknowledgements

This work was funded by the German Federal Ministry of Education and Research (reference number 01GQ0443).

References

[1] MJ Ferragamo and D Oertel. *J. Neurophysiol.* **87** (2002) 2262–2270.

[2] M Holmberg and W Hemmert. *Proc. CFA/DAGA'04* (2004) 773-4.

[3] JS Rothman and PB Manis. *J. Neurophys.* **89** (2003) 3097–3113.

[4] G Svirskis, V Kotak, DH Sanes and J Rinzel. *J. Neurophysiol.* **91** (2004) 2465–2473.