A MODEL FOR FILTERING AND ANALOG-TO-PULSE CONVERSION ON THE PERIPHERY OF AUDITORY PATHWAY

N.G. BIBIKOV, N.A. DUBROVSKY, G.A. IVANITSKY, L.K. RIMSKAYA-KORSAKOVA, V.N. TELEPNEV

N.N. Andreev Acoustical Institute, Moscow, USSR, 117836

ABSTRACT

An approach based on modelling the significant features of auditory processing, according to physiological evidence, provides a perspective opening into future speech analysis techniques. A model for information processing on the periphery of auditory pathway is presented. The model accommodates our knowledge of auditory nerve impulsion patterns and comprises a basilar membrane filtering, signal integration and rectification, and an analog-to-pulse conversion in first order auditory neurons which possesses refractoriness. Inserting the accumulating refractoriness in the model for auditory nerve fiber provides realistic representation of the short-term adaptation phenomena. The patterns of model reaction to tone bursts and stationary fragments of vowels are given.

The peripheral components of our model are represented by a bank of linear bandpass filters for simulation of the frequency selectivity of the inner ear [1], and a receptor-neuron model is to simulate generation of neural discharges. The impulse response of the minimum phase filter is as follows

\[ h(t) = e^{\omega \tau} \sin(\omega t) \]

Here \( \omega \) is the central frequency, and parameters \( \lambda \) and \( \beta \) define steepness of the front and rear slopes of impulse response. The specific values of parameters \( \lambda \) and \( \beta \) were chosen according to known data on auditory pathway periphery impulse responses. Both data on basilar membrane response to short acoustical clicks [2] and data obtained by reverse-correlation technique [3,4] were used. The reverse correlation between spike activity of auditory nerve fiber and the noise stimulus at input yields an estimate of impulse response. The result of quality of fillers used is not dependent on central frequency being approximately 4.2 if measured for 10 dB SPL. We avoided the modelling of compressive non-linearity of basilar membrane mechanical oscillations and limited ourselves with studies of relatively weak signals. In order to take into account signal transformations in hair cells, a partial one-period rectification was included in the model. Amplitude of signal's negative half-wave was three times less than the positive half-wave, in accordance with physiological data of Sellick and Russel [5]. After rectification and filtering the signal was added to Gaussian noise which provided random fluctuations of neuron membrane potential and presence of auditory nerve fibers spontaneous activity. A sum of noise and determinated signal then passed through the low-pass filter with integration time-constant of 8.2 as that corresponds to the known data on integrating properties of hair cell - spiral ganglion neuron dendrite system [6], and then fed the threshold circuit. We omitted modelling the phenomena of first synapse neuro-transmitter depletion since we could not find direct physiological evidence of this effect.

We paid much attention to modelling the postspike changes in the auditory nerve fiber. This changes, we consider, may play a significant role in coding variations in signal amplitude. We tried to evaluate refractoriness parameters from data obtained by Gaumond et al. [7]. These authors obtained hazard functions for auditory nerve fibers spontaneous activity using large statistical selections.

We succeeded in making our model reproduce these
functions. The best agreement with experimental data was achieved when the model parameters were as follows: standard deviation of intrinsic noise - 0.8 of threshold at rest; absolute refractoriness - 85 ms; relative refractoriness comprises two "phases" - "fast" and "slow" with the values of time constants 8.35 ms and 20 ms respectively (Fig. 1) [8]. The model, with parameters set to these values, was then exposed to high-frequency tone bursts. These bursts, after rectification and smoothing, become nearly rectangular depolarizing pulses of 58 ms duration each. We were interested in finding out whether or not such a model will show the intrinsic short-term adaptation properties, typical to the auditory nerve fibers [9]. It reveals that the shape of the model post-stimulus-time histograms depends crucially on the mode of summation of threshold elevations: namely, whether or not successive threshold elevations (relative to 20 ms "slow phase" refractoriness) are accumulated. If, at the end of absolute refractoriness, the threshold rises up to some fixed level, adaptation reveals to be very short and insignificant. But if, at the end of absolute refractoriness, the threshold increases for certain value beside the value it has had just before the spike occurrence, situation changes radically. In this latter case of accumulating refractory threshold changes the model shows the intrinsic property of short-term adaptation and decrease in ability for excitation succeeding the end of stimulus (Fig. 2).

Unfortunately we know few about the auditory nerve refractory patterns. There are data obtained by Sou bard et al. [19] showing successive inter-spike interval negative correlation; this corresponds to accumulating refractoriness hypothesis. Some additional data were published by Lutkenhouse and Smith [18]. A simple method of detecting the presence and measuring the time-course of accumulating refractoriness is provided by a comparison of two functions: the autocorrelation of stationary spike train, and the autocorrelation derived from interspike interval distribution obtained from the same spike sequence. If there is no accumulating refractoriness, and hence no inter-spike intervals correlation, these functions are identical, then accumulation is present, autocorrelation function drives lower than autocorrelation function and the difference between these functions represents the time course of accumulating refractoriness (Fig. 3).

The model that includes all stages mentioned above was studied respective to some vowel-like signals. Fig. 4 shows behavior of the model with total of 16 frequency channels that cover 110 - 800 Hz band. There can be seen 12the ability to distinguish formants, 2the representation of main tone with appropriate segregation; 3line dispersion of low frequency components; 4distinction of waveforms fine temporal peculiarities.

We do understand that the suggested model represents auditory periphery's properties roughly enough. Further development of the model implies involving non-linear smail signal amplification and large signal compression mechanisms. The model, as modified, would become free of dynamic range restrictions. Our techniques may have much more of a physiological and speculative motivation than a mathematical one. But we hope that involving such approach may become useful when designing speech recognition systems with neuronal interactions.
REFERENCES