THE MATLAB IMPLEMENTATION OF BINAURAL PROCESSING MODEL SIMULATING LATERAL POSITION OF TONES WITH INTERAURAL TIME DIFFERENCES

J. Bouše, V. Vencovský

Department of Radioelectronics, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

Abstract

The implementation of binaural auditory model able to reflect the lateral position of tones with interaural time differences is presented. The model is composed of two parts, monaural processing model adapted from Dau [2] and the binaural processing model designed by authors. The binaural processing model is simulating medial superior olive (MSO), part of human brain stem which is claimed to be responsible for coding of the temporal differences between signals in two ears [1]. Designed model is not using most widely used framework for binaural models, Jeffress' delay line, instead of it own designed approach inspired by Grothe's paper [3] is implemented. The output of the model was compared with the subjective listening tests taken from literature [8]. The results show that presented model is able to reliably reflect the subjective data.

1 Introduction

The human hearing system allows us to localize sound source in space. This ability is connected with the presence of two ears on our head. Since each ear lies on the contralateral side of the head the signals at their inputs coming from and outside sound source generally differ in time (figure 1) and intensity. These cues are called interaural level differences (ILD) and interaural time differences (ITD). The decoding of spatial information based on the ITD is believed to be processed in the first joint of ipsilateral and contralateral nerve path in the medial superior olive (MSO) a part of the human olivary complex in the brainstem [1]. During laboratory measurements when the subject is listening to the signal with the ITD through the headphones, the sound image is perceived within the head. Increasing the ITD will cause the shifting of the sound image intercranialy towards one ear. This phenomena is called lateralization [1].

This paper describes the implementation of the MSO model simulating the perception of lateralization caused by the ITD when listening through the headphones. A lot of different binaural models able to detect ITD were developed during last years [7]. Most of them is using as a framework the delay line proposed by Jeffress [5]. There is no evidence for existence of such structure in the mammalian olivary complex. Due this fact in proposed model the Jeffress delay line is not used instead of that the realization inspired by the neurophysiological findings described in Grothe's paper [3] is used. Grothe in cited paper found proofs of inhibition synapses found in the mammalian MSO coming from the contralateral ear through the auditory nerves with very short transmission time. It means that spikes transferred by these nerves should reach the MSO earlier than neural signals coming from both ipsilateral and contralateral ear to the excitatory inputs of the MSO. He claimed that these contralateral time shifted innhibition inputs are essential to detect ITD by the mammalian auditory system.
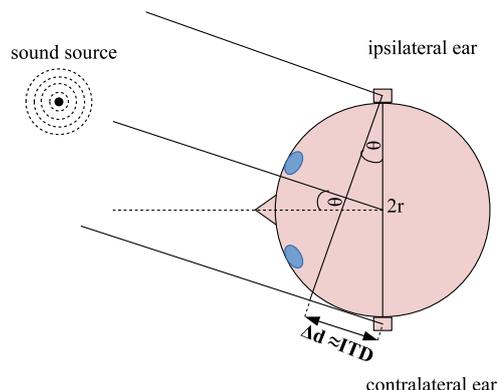

Figure 1: The differences in sound propagation paths resulting in the ITD

2 Model

The proposed model is composed of two main stages (figure 2). The first is monaural auditory model adapted from Dau's paper [2]. The second part is binaural model designed and implemented by the authors of the paper.
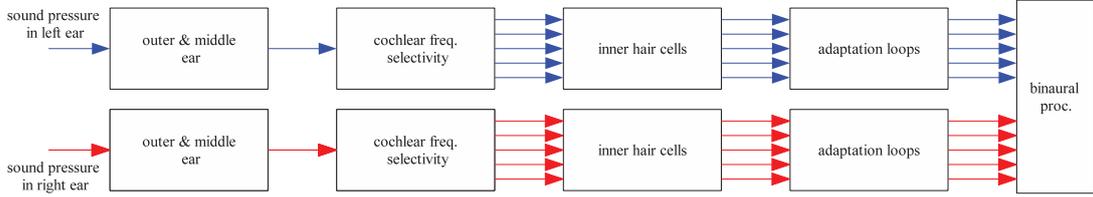

Figure 2: The model diagram

2.1 Monaural processing stage

The monaural model represents the input for the analyzed binaural signal. Since humans have two ears, there are two monaural parts in the overall model, each of them representing one ear. This model consists of the outer- and middle-ear model, cochlear frequency selectivity model, inner hair cell model and adaptation loops. The model was taken from Dau's paper [2].

The outer- and middle ear are modeled as linear system with certain frequency response by a cascade of FIR filters (256, 265 and 512th order). The FIR filters were designed in order to reflect experimentally measured amplitude transfer function of the outer- and middle-ear taken from the papers [6, 4].

The second part of the monaural processing stage is a cochlear selectivity model. This stage consists of the 4th grade gammatone filterbank which divides the input signal into 36 frequency bands according to equivalent rectangular bandwidth (ERB) function. Bandwidth of each gammatone filters is set to be equal to ERB value of the corresponding center frequency [2]. The minimal and maximal frequency of the filter bank were chosen to be 100 Hz and 15.5 kHz.

Signal in each band is then processed by the inner hair cell model and adaptation loops. The inner hair cells model consist of half-wave rectification followed by the 1st-order, low-pass IIR filter with cut off frequency of 1 kHz. This process roughly simulates the transformation of the basilar membrane vibration into the membrane potential inside the inner hair cell. If the output signal of this stage is lower than certain threshold value, it is then replaced by this threshold value. This simulates the absolute threshold of the human auditory system.The next part of the model are the adaptation feedback loops. Five consecutive, non-linear, feedback loops with the time constants of 5 to 500 ms model the temporal masking and adaptation. Input-output function of the loops approximates a logarithmic compression. It is non-linear for stationary input signals and it shows enhancement for rapid signal fluctuations (faster then the time constants in the feedback loops) [2].

2.2 Binaural processing stage

The binaural processing stage diagram can be seen in figure 3. Input signals from both ears are first filtered by first order IIR filter ($f_c = 300$ Hz). This roughly simulates lost of synchronization of the nerve cells on higher frequencies. It is assumed that output of the MSO is not affected by the level or difference in level of the input signals [1]. To eliminate the influence of the signal level is the signal normalized right after the half-wave rectification. The normalization is done
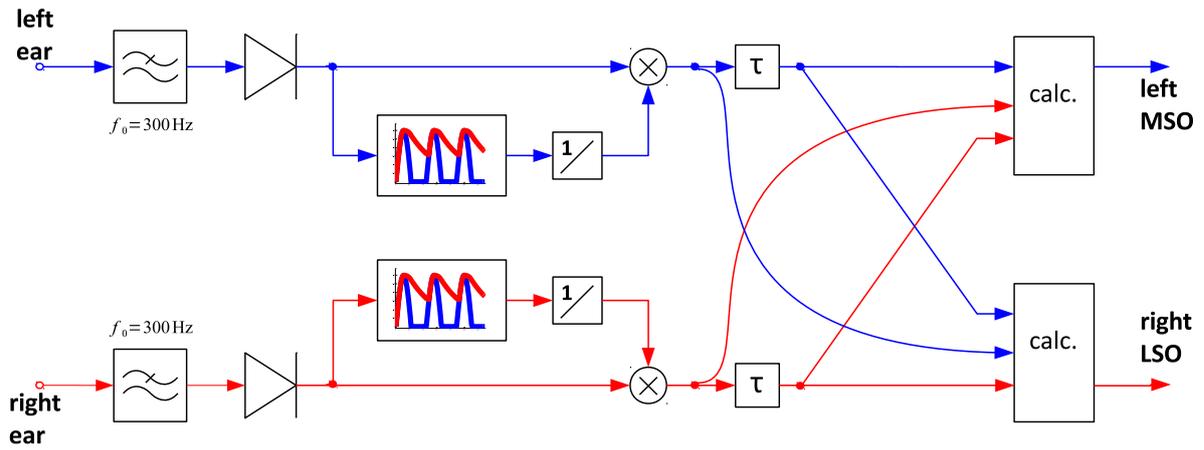
Figure 3: The binaural processing diagram

by dividing the input signal by it's envelope, which is extracted from the signal by the equation:

$$env(n, b) = \max(filt(n, b); orig(n, b)), \tag{1}$$

where env is envelope of the signal, $filt$ is filtered input signal by first order IIR filter with experimentally set time constant equal to 5 ms, $orig$ is input signal, n is sample number and b is number of ERB channel. The extraction of the envelope is denoted on figure 4.
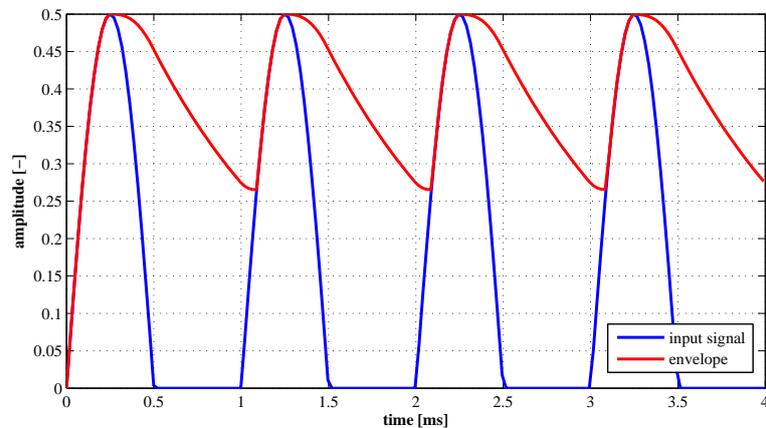

Figure 4: The extraction of envelope from the signal.

The signals from both ears are then processed in two calculation units, each of them with 3 inputs. Two of the inputs are the delayed signals representing the signals from ipsilateral and contralateral ear coming to the MSO. The delay was experimentally set to 180 μs. Third input is not delayed and represents the inhibitory signal from the contralateral ear. This design was inspired by Grothe's paper [3].

In the calculation units following mathematical processing is done

$$MSO(n, b) = Ip(n, b) \cdot Con(n, b) - CoInh(n, b) \cdot Ip(n, b) \cdot Con(n, b), \tag{2}$$

where Ip and Con is delayed signal from ipsilateral and contralateral ear respectively, $CoInh$ is contralateral inhibitory signal without delay, n is sample number and b is number of ERB channel. The MSO signal from both hemispheres is then again half-wave rectified and the average value in each of 36 channels is computed.

The signal is then processed by a cognitive model in order to obtain data comparable to the subjective lateralization experiments.

$$L(b) = \begin{cases} -1 \cdot (1 - r1(b)), & \text{left ear is leading} \\ (1 - r2(b)), & \text{right ear is leading,} \end{cases} \tag{3}$$

where:

$$r_1(b) = \frac{MSO_l(b)}{MSO_r(b)}, \tag{4}$$

$$r_2(b) = \frac{MSO_r(b)}{MSO_l(b)}, \tag{5}$$

MSO_r and MSO_l is the averaged signal from right and left ear respectively. The obtained scale ranges from 0 to 1, where 0 represents the perception of sound in the middle of the head and 1 near to the analyzed ear. This processing is denoted on figure 5.
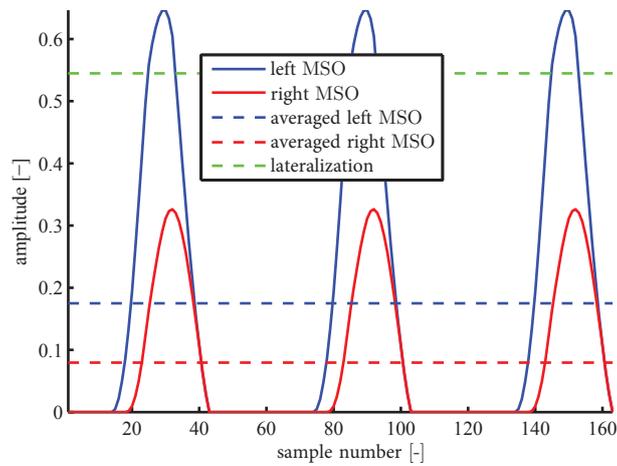

Figure 5: The output from MSO calculation units together with lateralization for phase shift equal to 45°
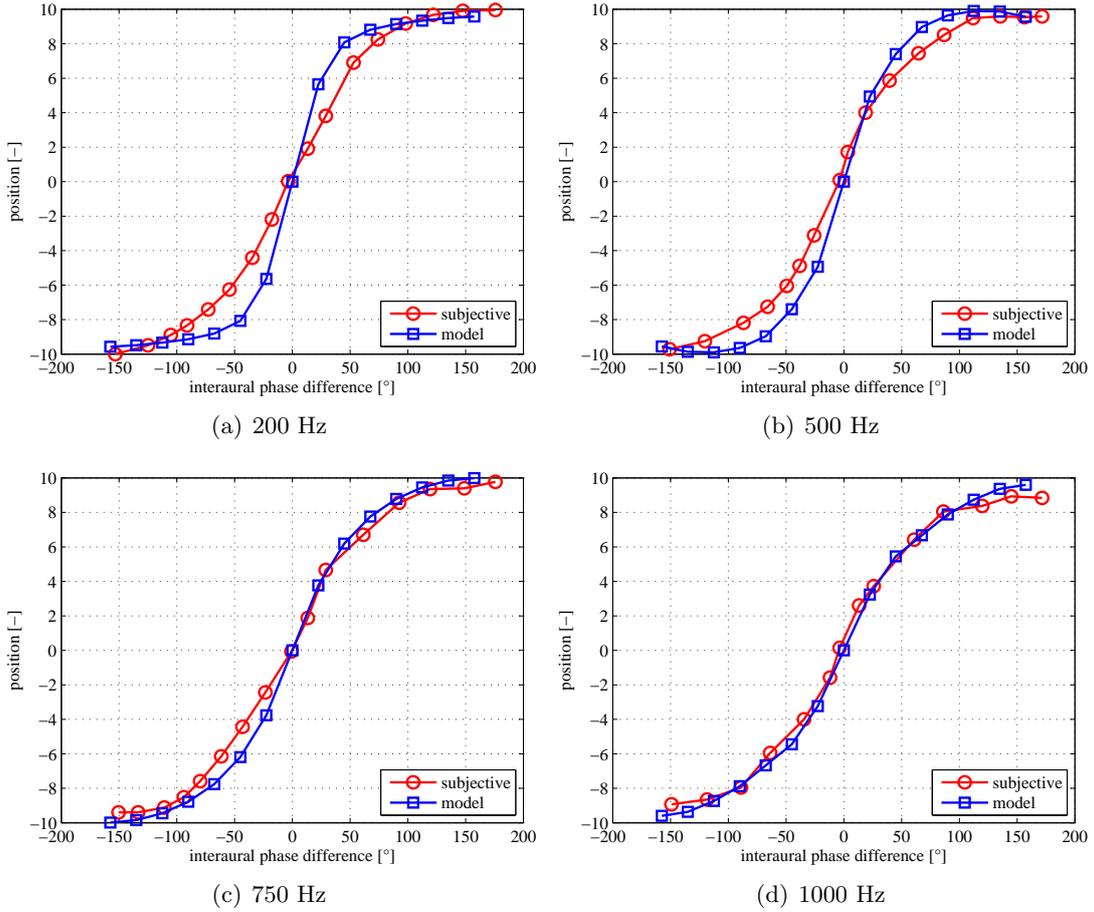
3 Results

Figure 6: Lateral position of auditory event as a function of Interaural Phase Difference of tone. Red line represents psychophysical data [8], blue line is data obtained by the auditory model.

Since a lot of lateralization experiments using signals with different ITD were already done and presented in the literature by other authors (see [1]), they were taken in this paper as a comparison to the modelled results. The data obtained by Yost [8] who measured lateralization of interaurally phase shifted sinusoids were particularly used. The sinusoids were interaurally shifted from -180 to 180 degrees.

The results were obtained for 50 dB SPL pure tones of four different frequencies 200 Hz, 500 Hz, 750 Hz and 1 kHz. They can be seen on figure 6. Just mean values of the psychophysical data are shown. Since their variance is quite high, modelled data are in all cases inside this variance.

4 Conclusion

The binaural auditory model suitable for simulating lateral position of sound sources via ITD was designed and implemented in Matlab. The model was able to simulate the psychophysical lateralization experiment with tones. In contrast to the most often used binaural models, this model does not use Jeffresss delay line [5]. Instead of that, time shifted contralateral inhibition theory presented by Grothe [3] is applied. Presented results shows good agreement between simulations and subjective experiments. Presented model thus can serve as a proof for the Grothes paper [3]. Its advantage in comparison to models using Jeffresss delay line is also in lower demands on the computation power.

Acknowledgement

References

[1] Jens Blauert and John S. Allen. *Spatial Hearing - The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, rev. edition, 1997.

[2] Torsten Dau, Birger Kollmeier, and Armin Kohlrausch. A quantitative model of the effective signal processing in the auditory system. i. model structure. *Acoustical Society of America*, 99:3615–3622, 1996.

[3] B. Grothe. New roles for synaptic inhibition in sound localization. *Nature Reviews Neuroscience*, 4(7):540–550, July 2003.

[4] D Hammershø I and H Mø ller. Sound transmission to and within the human ear canal. *Acoustical Society of America*, 100:408–427, July 1996.

[5] Lloyd A. Jeffress. A place theory of sound localization. In *Journal of Comparative and Physiological Psychology* [7], pages 35–39.

[6] M Kringlebotn and T Gundersen. Frequency characteristics of the middle ear. *Acoustical Society of America*, 77:156–164.

[7] R. Meddis. *Computational Models of the Auditory System*. Springer Handbook of Auditory Research, 35. Springer US, 2010.

[8] W. A. Yost. Lateral position of sinusoids presented with interaural intensive and temporal differences. *Journal of the Acoustical Society of America*, 70(2):337–409, 1981.

Jaroslav Bouše
Katedra radioelektroniky, FEL ČVUT v Praze, Technická 2, 166 27, Praha 6
tel. 22435 2109, e-mail: `bousejar@fel.cvut.cz`

Václav Vencovský
Katedra radioelektroniky, FEL ČVUT v Praze, Technická 2, 166 27, Praha 6
tel. 22435 2109, e-mail: `vencovac@fel.cvut.cz`