Phase difference representation of interaural timing disparities

Mathias Dietz, Stephan Ewert, Volker Hohmann
Medizinische Physik, Universität Oldenburg, 26111 Oldenburg, Deutschland, Email: mathias.dietz@uni-oldenburg.de

Introduction

The intracranial locus of a sound source is dependent on both interaural timing- and interaural level differences. A well-established model of interaural timing perception is the so called "Jeffress-Model" [1], which consists of several coincidence detectors along an internal delay-line. The interaural time difference (ITD) generated by a dichotic signal is estimated by the position of the highest coincidence in the delay-line. A modern variant of this model is based on contralateral inhibition (EI-model) [2]. Instead of using short term correlation for coincidence detection, the EI-model subtracts the differently delayed stimuli and uses the reduction of neural firing as a cue. Physiological findings in the barn-owl have revealed that this species uses a delay-line [3]. However, physiological investigations of guinea pigs do not unambiguously support the implementation of delay-lines in mammals [4]. An alternative concept is rate coding [5], in which the firing rate of a neuron is a monotonous function of the signal ITD. Hence, the ITD can unambiguously be determined by the given firing rate. Due to the quasi-periodic structure of the signal after bandpass filtering in the auditory periphery, the requirement of injectivity cannot hold for a long interval on the ITD axis. The investigations of Brand et al. [6] reveal that the point at which the firing rates reach their maxima and start to decrease are not a constant ITD but can roughly be described with $0.25/bf$, where bf is the best frequency of the respective neuron. If the ITD is substituted by the interaural phase difference (IPD) the firing rate maximum can be approximated as an independent value of $\pi/2$.

Given the physiological evidence that a delay-line might not exist in mammals, work into modeling interaural timing based on an IPD rate coding seems indicated. In this contribution we therefore present an effective binaural processing model which is based on IPD rate coding and does not need an internal delay-line. An information-theoretic approach is adopted in order to process IPD's.

A binaural processing model

The first stage of the model is the peripheral preprocessing as it is widely used for effective modeling (e.g. [2]). Outer- and middle ear are modeled with a 1^{st} order 1-4 KHz bandpass. The basilar membrane is modeled with a linear 4^{th} order gammatone filterbank [8]. For modeling the haircells a half wave rectification with a successive 770 Hz 5^{th} order lowpass filter is employed [2]. The model uses a static compression with a power of 0.4 (e.g. [9]). By adding white noise to all filterbands after the haircell transformation a finite hearing threshold is established.

In order to determine interaural phase differences separately for both the envelope (modulation) and the carrier (fine structure) two complex-valued second-order gammatone filters [8] are employed in parallel at the end of the monaural peripheral stage. One filter is centered around the same frequency as the first gammatone filter of the respective channel and therefore extracts the fine structure. Its bandwidth is set to the half of the respective center frequency. The second filter is centered at 150 Hz for all channels and is 150 Hz wide. This filter extracts a large range of modulation frequencies.

Since the outputs $g(t)$ of all filters, are complex-valued the phase $\phi(t)$ is given explicitly in the polar representation:

$$g(t) = a(t) \cdot e^{i\phi(t)}, \qquad (1)$$

where $a(t)$ is the amplitude of the signal $g(t)$.

Hence, the binaural processor does only need to subtract the phases of the left and right fine structure and modulation channels. We derive the IPD from the argument of the complex interaural transfer function (ITF) which can easily be derived in two different ways:

$$\mathrm{ITF}_1(t) = g_\mathrm{l}(t)/g_\mathrm{r}(t) \qquad (2)$$
$$\mathrm{ITF}_2(t) = g_\mathrm{l}(t) \cdot g_\mathrm{r}^*(t) \,. \qquad (3)$$

Finally, the ITF is 1^{st} order lowpass filtered at 64 Hz, in order to simulate a finite temporal resolution of the binaural processor [10]. The IPD(t) is now given as the argument of the filtered ITF. An interaural level difference can be extracted from the amplitude of the ITF, but only if it is derived as shown in equation (2).

For the following analysis, we assume that the auditory system can exploit both stationary and time-varying features that are preserved in the IPD(t).

Detection of temporal IPD fluctuations

In order to determine the temporal resolution of the model, perceptual data from broadband binaural beat (phasewarp) detection are investigated in the following. Fig. 1 shows the IPD as a function of time for a phasewarp stimulus with a binaural beat frequency of 4 Hz. The figure reveals that the model can easily resolve the beating.

Our analysis is based on the perception of phasewarp stimuli restricted to a frequency range of 1-550 Hz. Three interval, three alternative forced choice psychoacoustic measurements were conducted with four normal hearing listeners. The task was to distinguish the phasewarping from binaurally uncorrelated noise at 65 dB SPL and 500 ms interval length. The adaptive parameter was the

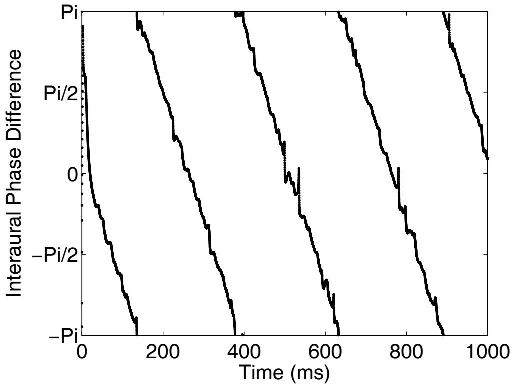

Figure 1: Display of a phasewarp IPD(t) extracted from the 500 Hz fine structure bands. The beat frequency is 4 Hz.

beat frequency. In condition 1, only frequencies from 1-550 Hz were presented to the listeners. In condition 2, binaurally uncorrelated noise was added in the frequency range 550-24000 Hz. The threshold beat frequency at which the listeners chose the correct interval with a 70% probability, was (77±13) Hz for condition 1 and (77±13) Hz for condition 2.

In order to compare these results with the IPD-model, Fig. 2 displays a 100-ms model output for 50 and 100 Hz. The five periods of the sawtooth-shaped 50 Hz beat

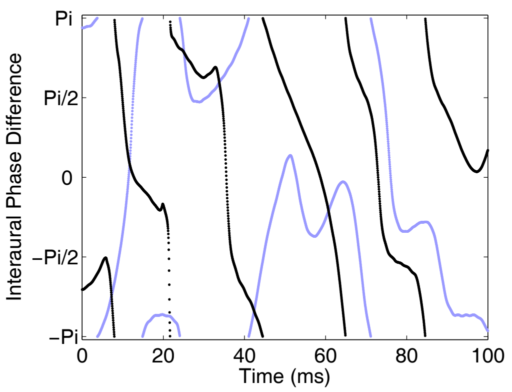

Figure 2: IPD as a function of time extracted from the 500-Hz fine structure bands for binaural beat frequencies of 50 Hz (black line) and 100 Hz (gray/blue line).

are clearly visible, but for the 100 Hz condition no periodicity can be observed. We want to emphasis that this effect is not due to the 64 Hz lowpass filtering. The filter frequency can be changed from 40 Hz to infinity without changing these displays significantly. However, the threshold beat frequency is always close to the bandwidth of the widest gammatone frequency band within the phasewarp range. Our model predictions are therefore in line with the psychoacousic measurement. However, they seem to contradict the model results from Breebaart et al. [11]. Their EI-model uses a double exponential smoothing window with a time constant of 30 ms, motivated by masking experiments. With this approach a 50-Hz binaural beat frequency can not be tracked. Tracking of beats up to about 75 Hz can be achieved by reducing the time constant of the EI-model to 1 ms.

Conclusion

Employing a model based on interaural phase differences interaural timing disparities can be extracted without a delay-line. The representation of stationary and fluctuating IPDs is in line with psychoacoustic findings and with simulations of an EI-model [2, 11] where the smoothing time constant is reduced significantly.

References

[1] Jeffress, L.A.: A place theory of sound localization. Journal of Comparative and Physiological Psychology 41 (1948), 35-39

[2] Breebaart,J., van de Par, S., Kohlrausch, A.: Binaural processing model based on contralateral inhibition. I. Model structure. Journal of the Acoustical Society of America 110 (2001), 1074-1088

[3] Carr, C.E. and Konishi, M.: A Circuit for Detection of Interaural Time Differences in the Brain Stem of the Barn Owl. Journal of Neuroscience 10 (1990) 3227-3246

[4] Mc Alpine, D., Jiang, D. and Palmer, A. R.: A neural code for low-frequency sound localization in mammals. Nature Neuroscience 4 (2001), 396-401

[5] Mc Alpine, D. and Grothe, B.: Sound localization and delay lines – do mammals fit the model? Trends in Neurosciences 26 (2003) 347-350

[6] Brand, A., Behrend, O., Marquardt, T., McAlpine, D. and Grothe, B.: Precise inhibition is essential for microsecond interaural time difference coding. Nature 417 (2002) 543-547

[7] Van de Par, S. and Kohlrausch, A.: A new approach to comparing binaural masking level differences at low and high frequencies. Journal of the Acoustical Society of America 101 (1997), 1671-1680

[8] Hohmann, V.: Frequency analysis and synthesis using a Gammatone filterbank. Acta Acustica united with Acustica 88 (2002), 433-442

[9] Ewert, S.D. and Dau, T.: The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. Journal of the Acoustical Society of America 108 (2000), 723-734

[10] Siveke I., Ewert, S., Wiegrebe L.: Perceptual and physiological characteristics of binaural sluggishness. International Symposium on Hearing (2006), 354-357

[11] Breebaart,J., van de Par, S., Kohlrausch, A.: Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters. Journal of the Acoustical Society of America 110 (2001), 1075-1104