



19th INTERNATIONAL CONGRESS ON ACOUSTICS MADRID, 2-7 SEPTEMBER 2007

A CLOSER LOOK AT THE REPRESENTATION OF INTERAURAL DIFFERENCES IN A BINAURAL MODEL

PACS: PACS: 43.66 Pn

Nicolas Le Goff¹ ; Armin Kohlrausch² ¹ ; Jeroen Breebaart²

¹ Technische Universiteit Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands;
n.legoff@tm.tue.nl

² Philips Research Europe, High Tech Campus, 5656 AE Eindhoven, The Netherlands;

ABSTRACT

In this contribution, we investigate the internal representations given by the binaural model proposed by Breebaart et al. [1]. We focus on the cues used by an artificial observer when determining just noticeable differences in interaural time differences (ITDs) or interaural level differences (ILDs). The binaural processor consists of an array of excitation-inhibition (EI) cells, each characterized by its internal interaural delay and internal interaural attenuation. This model is conceptually similar to a cross-correlation model and the relevant properties for ITD and ILD discrimination tasks are changes in the binaural pattern along the internal delay and the internal attenuation axis. For ITD discrimination of two coherent signals, the traditional approaches consist of searching for a displacement of the peak of the normalized cross-correlation function or measuring the difference in the cross-correlation function at lag zero. These approaches will be compared to our strategy which maximizes the differences between the two normalized cross-correlation functions. Threshold predictions depend on which position along the internal delay line is analyzed. The characteristics of the binaural pattern with regard to ILD discrimination tasks will also be investigated and compared to those that are relevant for ITD discrimination.

INTRODUCTION

The human auditory system senses its environment with two ears, as such, it is qualified as a binaural system. While this is an obvious fact, the way our binaural auditory system is working remains a matter of investigation. The human auditory system has such a degree of complexity that it is traditionally divided into several parts that are studied in isolation. In an attempt to extend knowledge about modeling of the human auditory system, we focus our research toward binaural modeling. In this context, we propose to have a closer look at the internal representations given by a binaural model which has the property to include both peripheral and binaural processing. This type of investigation is made possible by the model proposed by Breebaart et al. [1]. This model was originally designed to simulate binaural masking conditions. We are now interested in bringing the model to a new area, namely ITD and ILD detection. To do so we need to understand what are the characteristics and the nature of the cues that are usable for these tasks. We therefore study the characteristics of the signals at the output of a binaural processor.

A BINAURAL MODEL

The signal processing model that we use in this study was developed by Breebaart et al. It extends the monaural model of Dau et al. [2] with a binaural processor and an optimal detector that operates on the output of the binaural processor, as well as monaural inputs. It is in essence similar to an equalization-cancellation model. In this model the peripheral processing is preceding the binaural processing and it is fundamental to consider both processes at the same time. An introduction to the relevant processes in the model for the present study is given. Further details

can be found in [1]. A simplified representation of the model with the relevant parts for this study is shown in Fig.1. The peripheral part of the hearing pathway is simulated through different stages.

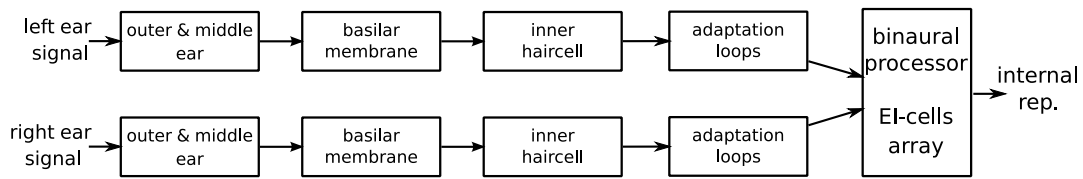


Figure 1: A simplified scheme of the binaural model.

The first stage combines an outer and middle ear filter that consists in this study of a bandpass filter that is sufficient to simulate headphone conditions. The second block simulates the basilar membrane and is modeled by a linear third-order gammatone filter bank. In simulations used for determining absolute detection thresholds an internal noise is introduced at this stage (not shown in Fig.1). The third block represents the inner haircell processing and consists of a half-way rectifier followed by a fifth-order low-pass filter with a -3 dB cutoff frequency of 770 Hz. The last stage includes the influence of adaptation with various time constants. The output of this stage for a block function will first exhibit a large overshoot, then show a steady state behavior which will be followed by a decay with a modest undershoot. For a signal with a flat temporal envelope, the input-output characteristics in steady state is nearly logarithmic and the output is expressed in model units (mu). Both left and right ear signals go independently through the peripheral stages.

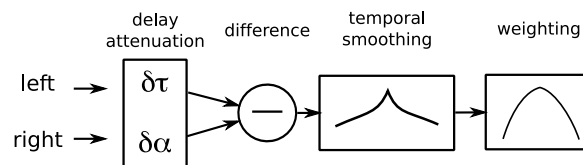


Figure 2: A simplified scheme of an EI cell

The output of the adaptation loops are fed to the binaural processor. The binaural processor consists of an array of excitation-inhibition (EI) cells. Each cell is characterized by its specific internal delay τ and specific attenuation α . As shown in Fig. 2 an EI cell is taking both right and left input signals on which the specific delay and attenuation of the cell are applied. Then the difference of the two signals is computed. The resulting signal is then temporally smoothed with a double sided exponential window with a time constant of 30 ms. Finally the signal is weighted to refer to the observation that the density of EI cells differs with internal delay. Furthermore in the context of this study we do not consider internal noise that is normally used to limit the accuracy of the model when used as an artificial observer.

In theory the binaural processor is a combination of a virtually infinite number of EI cells. The combination of the outputs of all EI cells constitutes the internal representation of the left and right ear signals, and is depicted as the binaural pattern. In practice only a finite amount of EI cells are used for simulations. Depending on the nature of the left and right ear signals the relevant parts of the binaural pattern are located at different positions and have different characteristics. In headphone reproduction, lateralization can be obtained by either delaying the signal on one side or by attenuating the signal on one side. The ability to perceive a change in lateralization is referred to as ITD and ILD detection and are the two conditions for which we are studying the properties of the binaural pattern.

ITD CASE

At first we have a look at the binaural pattern when a delay is introduced between the left and the right ear signals. Left and right ear signals will always have the same intensity we can restrict the binaural pattern, which has by nature two dimensions (α, τ) to a one dimensional pattern, the τ -axis.

nature of the binaural pattern

The sinusoidal signals used in this paper are 400 ms long, have a frequency of 838 Hz and an amplitude of 65 dB SPL. Furthermore all binaural patterns and signal waveforms are shown for an auditory filter centered at 838 Hz. The binaural patterns for this signal with zero interaural delay and with an ITD of 100 μ s are presented in the left panel of Fig. 3 at two instances of time. The two upper curves represent the binaural pattern of the centered (solid line) and the lateralized (dashed line) sinusoid at the instance 50 ms. The two lower curves represent the binaural pattern of the centered (dash-dotted line) and the lateralized (dotted line) sinusoid at the instance 300 ms. The sinusoid without internal delay has a representation centered at 0 ms on the τ -axis. At this internal delay the curves reach 0 mu which means that there is a perfect cancellation of the two signals in the binaural processor. On each side of the 0-ms point, the curves increase until a maximum is reached. The curve is symmetrical around 0 ms and will be periodic along the τ -axis. The internal representations of the 100- μ s lateralized sinus (dashed and dash-dotted lines)

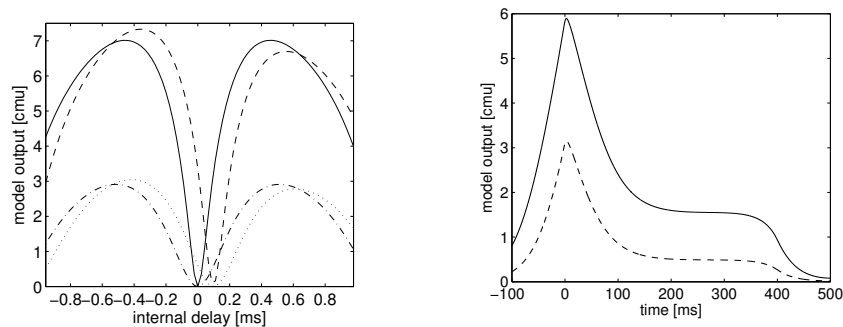


Figure 3: Left: binaural pattern of a 0 ITD and a 100 μ s ITD 838Hz sinusoid at instance 50 ms (two top curves) and at 300 ms. Right: waveform of the output of the -0.068 ms EI cell for the 100 μ s lateralized (solid line) and centered sinus (dashed line).

are somewhat similar. However the cancellation is not perfect and the minimum of the patterns does not reach its minimum for 0 ms but for an internal delay of 100 μ s. In the right panel, the output of the EI cell characterized by an internal delay of -0.068 ms for lateralized (solid line) and a centered (dashed line) sinus are plotted. An EI cell with a non-zero τ was selected to demonstrate the temporal behavior of the output. The presence of information for negative time and after end of the signal (time > 400 ms) is due to filtering in the peripheral stages and temporal smoothing in the binaural processor. For both signals shown in the right panel of Fig. 3, the waveform exhibits first a strong overshoot which is followed by a steady state where the output is stabilized until its decay.

optimizing the detection

One way to detect an ITD in the model is by looking at a displacement of the minimum of the EI pattern. Rather than looking at the 0 point along the τ -axis, previous studies have already reported the possible benefit of considering other positions (Colburn et al. in [3]). Therefore, we propose to perform the detection of an ITD by choosing the EI cell(s) that will maximize the difference between the binaural pattern of the centered version and the lateralized version of a signal. This concept is somewhat following the idea of a best delay place along an internal delay axis that has been previously mentioned in the literature (see [4]). Consequently we compute the difference between the two binaural patterns corresponding to the lateralized and the centered version of the signal. The resulting EI output difference as a function of τ is shown in the left panel of Fig. 4 for a given instance in time. The curve elicits two extrema. Both extrema are by definition the places along the τ -axis that maximize the difference between the two binaural patterns. These two extrema are referred to as optimal EI cells. The curve in the right panel represents the time course of the difference of the outputs of the binaural processor between the lateralized and centered signal taken at one optimal EI cell.

The difference of the binaural patterns in the left panel of Fig. 4 was computed for a single instance in time. It is possible to compute such a curve for each sample of a signal. For each

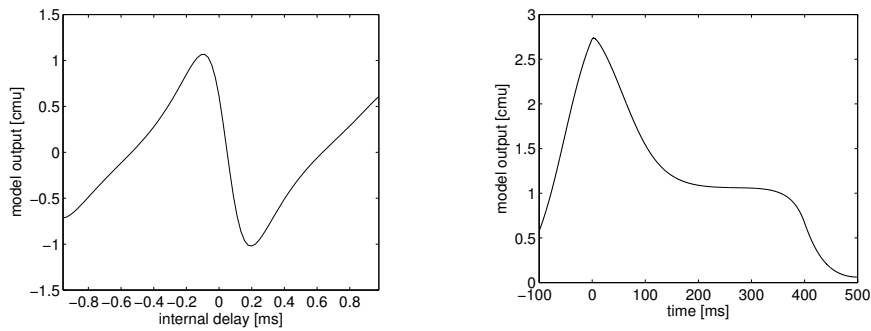


Figure 4: Left: Difference of binaural patterns between a 100 ms ITD and a 0 ITD sinusoid as a function of the internal delay. Right: difference of the output of the binaural processor for one optimal EI cell as a function of time.

sample the curve would be different but it is possible to extract valuable information of such a process. It is beyond the scope of this contribution but it can be shown that despite the presence of the overshoot, the positions of the optimal EI cells can be considered constant over the duration of the signal. It can also be shown that positions of the optimal EI cells do not directly depend on the amplitude of the signals. However the positions of the optimal EI cells depend on the external ITDs and it is possible to derive a relationship between the external ITD and the position of the optimal EI cells along the τ -axis. Consequently we have computed by model simulations, using sinusoidal input signals, the position of the two optimal EI cells for a range of useful external ITDs, and for individual auditory channels. The results of the simulations can be interpolated

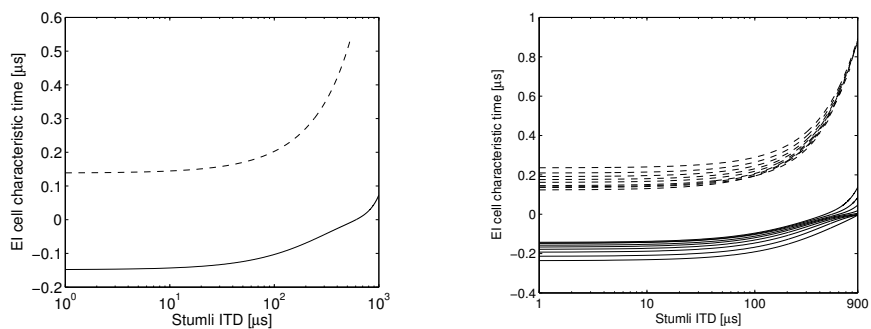


Figure 5: Left: ITD tuning curve for a single channel centered at 838 Hz. Dashed and solid lines represent for the two different optimal EI cells, Right: ITD tuning curves for 9 auditory channels between 320 Hz and 1100 Hz

and the resulting polynomials are called ITD tuning functions and their representations ITD tuning curves. The ITD tuning curves for nine auditory channels are shown in the right panel of Fig. 5. The dashed lines represent the positions of one of the optimal EI cells, the solid lines represent the positions of the other one. These curves show that the position of the optimal EI cells are depending on the considered auditory channel, each curve eliciting the same shape. In the left panel, the ITD tuning curve of the channel centered at 838 Hz in isolation shows that for small external ITDs ($<100\mu s$) the position of the optimal EI cell could be considered as constant. It also shows that for larger ITDs the position of the optimal EI cells are considerably shifted from their value at lower ITDs. These ITD tuning curves can now be incorporated in the binaural processor and detection thresholds for various types of conditions can be compared by using the so called optimal EI cells and other approaches.

ILD CASE & COMPARISON

In a similar manner, the investigations on the characteristics of the binaural pattern in the case of a lateralization created by attenuation of the signal on one side can be studied. The detection of such an attenuation is referred as ILD detection.

nature of the binaural pattern

As done previously, the binaural pattern for this part of the study can be restricted to one dimension, the internal attenuation axis (α -axis). The binaural pattern for a 400 ms, 838 Hz sinus presented in the left panel of Fig. 6 as a function of α at two instances in time (same convention of representation as for ITD). The solid and the dash-dotted line reaching the (0,0) point are the binaural pattern of the centered signal at 50 ms and at 300 ms. The temporal output of the binaural processor for an EI cell characterized by an internal attenuation of 2 dB and an input comprising a sinus with an ILD of 10 dB (solid line) and the centered sinus (dashed line) are plotted in the right panel. The waveforms of both signals exhibits a roughly similar shape as for the ITD case. A strong overshoot is first developed, then comes a steady state and finally the amplitude will drop (not shown). However there are also differences which will be discussed in the next section.

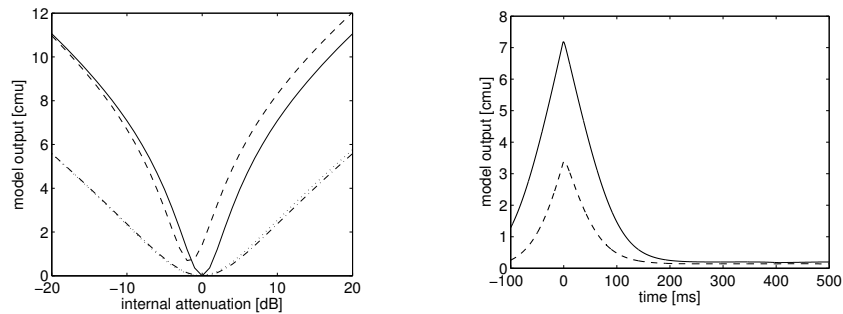


Figure 6: Left: binaural pattern of a 0 dB and a 10dB ILD 838 Hz sinusoid, at instants 50 ms (two top curves) and 300 ms. Right: waveform of the output of the 2 dB EI cell for the 10 dB lateralized (solid) sinus and the centered (dashed) sinus.

ILD vs ITD

Looking at the binaural patterns for ITD in the left panel of Fig. 3 and for ILD in the left panel of Fig. 6, allows us to make comparisons. In both cases, the patterns for the centered sinus are symmetrical around 0 dB or 0 ms regardless of the instant of time (solid and dash-dotted lines). While the binaural patterns of the ITD case are periodic, those of the ILD case are not. What happens in these patterns is similar to the properties of the cross-correlation of two signals where in one case one would be delayed and in another case one would be attenuated. In both cases, the binaural pattern of the lateralized sinus during the overshoot does not reach zero activity, which is a consequence of the decorrelation resulting from the behavior of the adaptation loops. While in the ITD case, the pattern during the overshoot is almost a scaled replica of the one in steady state. In the ILD case other non linearities are involved and we do not observe the same phenomenon. Furthermore the pattern of the lateralized signal for the ITD case is symmetrical around an internal delay equal to the external ITD i.e. 100 μ s. In the case of the ILD a very strong compression occurs and even though the pattern of the lateralized sinus is slightly shifted along the α -axis, the minimum of the curve is clearly away from a value corresponding to the external ILD, i.e. 10 dB.

Now looking in the time dimension rather than the internal delay or internal attenuation dimension, we can compare the waveforms of the outputs of the binaural processor shown in the right panel of Fig. 3 for the ITD case and in the right panel of Fig. 6 for the ILD case. The reader should remember that even though it is not shown here, the characteristics of these curves are greatly varying depending on the considered EI cell. The ones for which the output is depicted were chosen to provide a first insight. One can immediately notice that all curves present a strong overshoot. The steady state in the ILD case is however much lower than that for the ITD case. Furthermore the differences between the output of the centered and the lateralized signals are much smaller in the ILD case than the ITD one. The interesting point is that the characteristics of all curves are actually given by the same parts of the model. In the ILD case the difference of amplitude between the right and left signals will be emphasized by the adaptation loops during the overshoot thus creating the difference between the internal representation plotted in the right panel of Fig.

6. However, in steady state the logarithmic behavior of the adaptation loops will strongly decrease the difference between the two input signals and thus reduce the output to nearly zero for any EI cell. In the ITD case the situation is more complex. The presence of the overshoot, which is not as pronounced as for the ILD case, is also resulting from the adaptation loops but for different reasons. Adaptations loops are sensitive to the absolute start phase of the signals. Since one the signal is delayed, both left and right signals do not have the same absolute phase and will excite the adaptation loops in a different manner. The EI output in steady state, however is larger for the ITD case than for the ILD case because the considered EI cell along the τ -axis is away from zero. The even larger amplitude in steady state for the lateralized signal is mainly resulting from a combination of the subtraction in the binaural processor of the two phase shifted signals in combination with the temporal smoothing that occurs at the output of the binaural processor. These processes will result in the curves in the right panel in Fig. 4 for the ITD case and Fig. 7 for the ILD case. The two curves are not directly comparable in terms of amplitude because

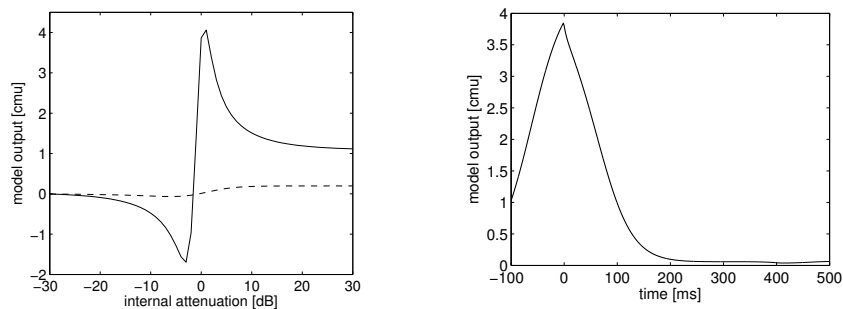


Figure 7: Left: Differences of binaural pattern between a 0 dB ILD and a 10 dB ILD for a 838Hz sinus. Right: Waveform of the difference of the output of the binaural processor for one EI cell.

an attenuation of 10 dB does not correspond to the same perceptual lateralization as a delay of $100\mu\text{s}$. From these representations we can see that, in the ILD case, the behavior of the model is totally determined by the behavior of the adaptation loops in their transient phase. In the ITD detection case, the contribution of the overshoot is less important though it remains a large part of the total information.

CONCLUSION

In this contribution we used a binaural model that can derive lateralization thresholds by considering the processes occurring in both peripheral areas and in the binaural processor. We proposed a strategy to mathematically describe the optimal places on the binaural pattern in order to perform ITD detection. Further investigations and simulation should be done to confirm the validity and the limitations of the tuning curves. Comparisons have been made between internal representations of stimuli for ITD and ILD cases and we have seen that two different techniques of lateralization, ITD and ILD, can lead to different characteristics of the binaural pattern.

References

- [1] J. Breebaart, S. van de Par, A. Kohlrausch. Binaural processing model based on contralateral inhibition I. model structure. *Journal of the Acoustical Society of America* **110** (2001) 1074–1088
- [2] T. Dau, B. Kollmeier, A. Kohlrausch. Modeling auditory processing of amplitude modulation. i. detection and masking with narrow band carriers. *Journal of the Acoustical Society of America* **102** (1997) 2892–2905
- [3] A. de Cheveigné, S. McAdams, L. Collet. *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*. Springer (2004)
- [4] L. A. Jeffress. A place theory of sound localization. *Journal of computational physiology and psychology* **41** (1948) 35–39