Audio Engineering Society

Conference Paper

Presented at the Conference on
Headphone Technology
2016 Aug 24–26, Aalborg, Denmark

The Effect of Elevation on ITD Symmetry

Andrea F. Genovese, Jordan Juras, Chris Miller, and Agnieszka Roginska

Music and Audio Research Lab - New York University

Correspondence should be addressed to Andrea Genovese (genovese@nyu.edu)

ABSTRACT

Typical HRIR modeling techniques ignore the issue of asymmetric head and pinnae characteristics among listeners, thus missing a possibly important correction parameter needed for realistic simulations. In fact, morphological asymmetry is very likely to be a cause of ITD asymmetry. In a previous exploratory study, a common region of sensitivity between datasets of individual HRIR measurement, where the ITD asymmetry is more prominent, was found in the azimuth range of $\pm 90°$ to $\pm 130°$ on the horizontal plane ($\phi = 0°$). This paper further expands the investigation of ITD asymmetry to the elevation dimension. Two widely used publicly available databases of individually measured HRIRs were selected and analyzed. Due to different measurement techniques and sample rates, the analysis was performed separately for each set. Results found that an increase or decrease in elevation angle ϕ would affect the sensitivity region by reducing the maximum and mean ITD asymmetry value in a roughly linear fashion. This fact implies that the impact of morphological asymmetry is gradually less severe as the elevation angle moves away from the horizontal plane.

1 Introduction

In Virtual Surround Sound, spatial audio is recreated by filtering mono or stereo sound files with special filter pairs called *Head Related Impulse Responses* (HRIRs). Each HRIR pair (left and right ear) is representative of a specific spherical location in reference to the listener, denoted by the azimuth angle θ and elevation angle ϕ.

By a convolution process between a mono-stereo audio file with an HRIR, the spatial cues contained within the HRIR are transferred to the audio file [1]. These cues are perceptually related to the spatial position for which the HRIR was recorded and are used by the human brain to decode a sound source's direction. This paper focuses on the ITD cue (*Interaural Time Delay*) which is the sound source's direct path time-arrival difference between the ipsilateral and the contralateral ear. ITDs are one of the most important spatial cues and the primary cue for low frequencies [2].

Due to the high variability of individual anthropometric characteristics among humans, generalized HRIRs recorded on plastic dummy heads do not necessarily represent perceptually viable spatial cues for listeners. It is recognized that a listener's localization accuracy is often determined by how closely his/her morphological parameters, i.e. the head and the ears, match those of the used dummy head. Furthermore, human heads and outer ears are seldom symmetric [3]. Asymmetry in pinnae and head shape might imply for some listeners a significant asymmetry in the spatial cues, in particular, ITDs. For these reasons, individualized HRIRs are considered to be desirable for a listener when experiencing binaural sound.

However since it is not feasible to measure HRIRs for the consumer market, modeling techniques seek to synthesize or adapt general HRIRs to a listener using his own anthropometric data, obtained, for example, through a photographic technique [4] [5]. Current modeling techniques often assume a symmetrical head, meaning that the spatial cues recreated for one of the hemispheres are also used for the opposite hemisphere (for example, the cues obtained for an azimuth angle $\theta = 30°$ are copied to the mirrored position $\theta = -30°$). There is an indication in literature that this assumption of symmetry does not necessarily reflect reality [6], and that to approximate the head to a symmetrical model might cause noticeable perceptual mismatches that could possibly affect localization performance in applications where accuracy is a key aspect.

A previous paper by the authors [7] showed the existence of relevant ITD asymmetry between the left and right hemispheres on the horizontal plane. A common region of sensitivity across the examined databases, where the asymmetry is more prominent, was found between $\theta \in [90°, 130°]$ for $\phi = 0°$. In this paper, the issue of ITD asymmetry is further extended to the elevation dimension. Two publicly available databases (also examined in [7]) of individual HRIR measurements are analyzed and inspected for elevation-dependance asymmetry between locations mirrored across the left and right hemispheres.

The motivation behind this analytical study is to further explore the severity of the effect of morphological asymmetry on personalized HRIR measurements. The issue of "perceptual asymmetry" [8] which validates whether the asymmetry is perceptually noticeable or not, is not covered by this paper which instead focuses on finding a statistical pattern in measured asymmetry for different azimuth angles at different elevations.

2 Databases

The data for this analysis was collected from publicly available databases of individualized HRIRs. The two selected databases (LISTEN [9], CIPIC[3]) differed in angle resolution and measurement technique. Due to these differences, it was decided to analyze each database independently instead of choosing a data pooling approach.

The sampling rate used would determine the temporal resolution of the detected ITDs in each HRIR pair. For a 44.1 kHz sampling rate, the distance between each sample, and, therefore, the minimum ITD value, is $22.6\mu s$.

2.1 LISTEN database

The HRIRs in this database were recorded by IRCAM and AKG for their collaborative LISTEN project [9]. HRIRs were recorded for variable azimuth resolutions (typically $\theta + 15°$ increments) at different elevations with $\phi \in [-45°, +90°]$ (15° increments). The measurements were conducted on 50 individual subjects for a total of 187 locations per subject at 44.1 kHz.

The measurement technique made use of a very precise crane structure connected to a loudspeaker rig and a software controlled rotating chair making the subject's azimuth alignment precision very reliable and less prone to error caused by human misalignment.

2.2 CIPIC database

UC Davis measured 45 subjects at 1250 spherical locations for 25 non-uniformly-spaced azimuth locations $\theta \in [-80°, +80°]$ and 50 elevations $\phi \in [-45°, +230.625°]$ (steps of 5.625°) [3]. Golay codes at 44.1 kHz were used to measure the HRIRs with the subject seated in a 1m radius hoop aligned on the subject's interaural axis. Subjects were not constrained but were able to monitor their head position.

The database includes complementary anthropometric data for 37 out of the 45 subjects. The anthropometric data consists in a detailed collection of measurements of head and pinnae parameters, for both hemispheres, that allow for possible useful correlational studies between ITD asymmetries and morphological asymmetries. Table 1 below illustrates a preliminary analysis on the symmetry of individual anthropometric data. The table reports the values of the average parameter difference across subjects (calculated as $|x_{left} - x_{right}|$) and its related standard deviation. The parameter that presented the most variability across subjects is the "Pinna Height", with 0.3207 cm average difference between left and right ears. It is not yet clear how severely these anthropometric asymmetries impact ITD asymmetry, but their presence supports the fact that mirrored ITD would also be asymmetric. The results obtained in [7] indeed show the presence of an asymmetry, indicating a likely correlation between the two.

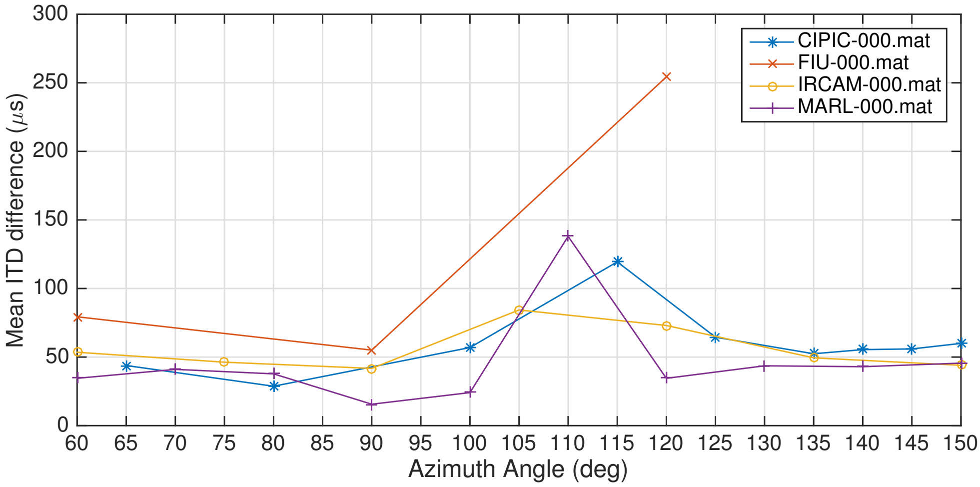

Fig. 1: Close-up superposition of average ITD asymmetry curves across 60° and 150° for the horizontal plane $\phi = 0$. An identified region of sensitivity spans between 90° and 130° (figure from [7])

Anthropometric Feature	Mean	σ
Cavum Concha Height (cm)	0.1213	0.1077
Cymba Concha Height (cm)	0.0910	0.0642
Cavum Concha Width (cm)	0.1357	0.0883
Fossa Height (cm)	0.1551	0.1421
Pinna Height (cm)	0.3207	0.2737
Pinna Width (cm)	0.1613	0.1251
Intertragal Incisure Width (cm)	0.0715	0.0551
Cavum Concha Depth (cm)	0.1295	0.1152
Pinna Rotation Angle (deg°)	0.0957	0.0415
Pinna Flare Angle (deg°)	0.0813	0.0638

Table 1: Mean and Standard deviation of the left-right difference of pinnae features measurements for the subjects in CIPIC

3 Horizontal Plane Asymmetry

The case of ITD asymmetry on the horizontal plane has already been explored in the authors' previous publication [7]. In that context, the ITDs were calculated as the distance in samples (then translated into seconds) between the two points of maximum time-domain cross-correlation between the left and right ear HRIR measurement. For consistency, the same ITD calculation method was applied also when the ITD data was already provided. In each dataset, the average asymmetry $\bar{S}(\theta)$ for $\phi = 0$ and subject set N was calculated using equation 1 as the average absolute difference of magnitudes between each ITD with its mirrored counterpart in the opposite hemisphere (for each available angle $\theta \in [0°, 180°]$ where a counterpart existed, thus excluding $\theta = 0°$ and $\theta = 180°$).

$$\bar{S}(\theta,\phi) = \frac{1}{N}\sum_{n=1}^{N} \left| |ITD_n(\theta,\phi)| - |ITD_n(360° - \theta,\phi)| \right| \tag{1}$$

Figure 1 shows the superposition of the average asymmetry curves for four different databases over a region spanning from 60° to 150°. The resulting graphs show an interesting common region of sensitivity roughly between 90° and 130° azimuth, where the curves are more prominent. The graph suggests that ITDs in this region are more susceptible to anthropometric asymmetries.

4 Analysis

Similarly to the previous study [7], the different azimuth and elevation resolutions found in the datasets,

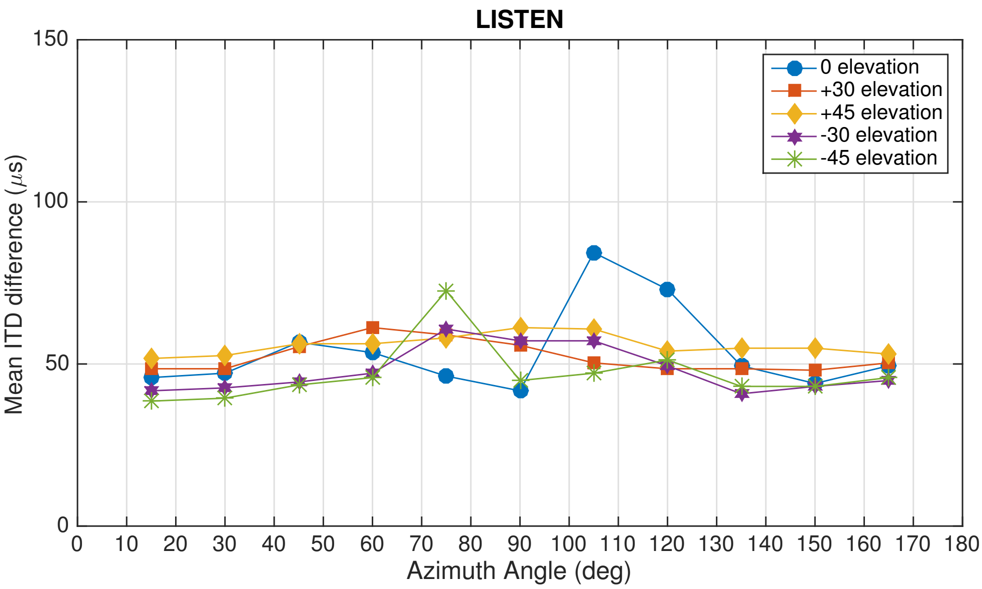

Fig. 2: Elevation curves for the LISTEN database

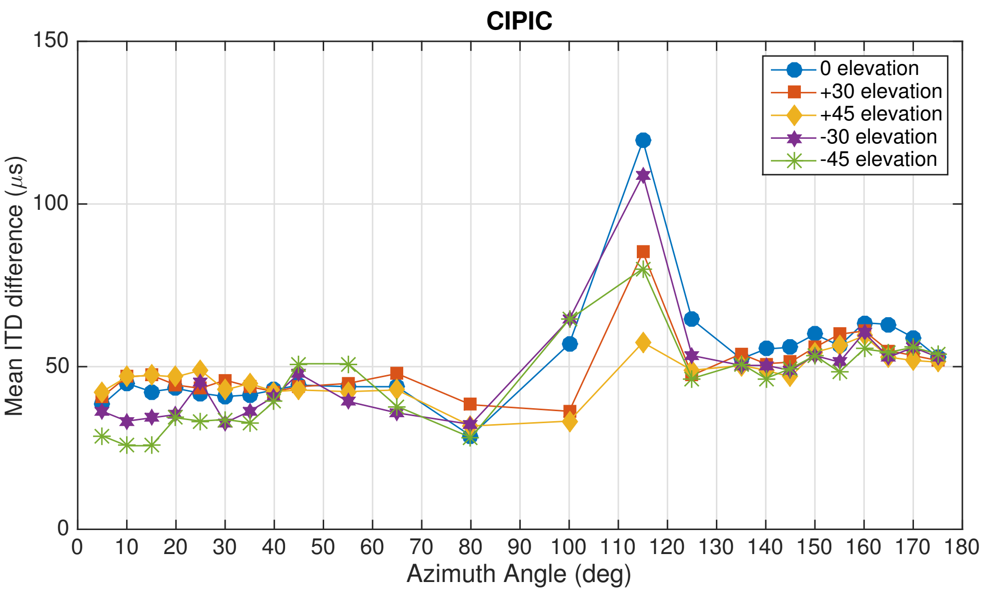

Fig. 3: Elevation curves for the CIPIC database

Database	LISTEN (N=50)				CIPIC (N=45)			
ϕ	$\theta_{\mathbf{peak}}$	$\bar{\mathbf{S}}(\theta_{\mathbf{peak}},\phi)$	μ	§.E.	$\theta_{\mathbf{peak}}$	$\bar{\mathbf{S}}(\theta_{\mathbf{peak}},\phi)$	μ	§.E.
-45	75°	$73\mu s$	$47\mu s$	$1\mu s$	115°	$80\mu s$	$45\mu s$	$2\mu s$
-30°	75°	$61\mu s$	$48\mu s$	$1\mu s$	115°	$108\mu s$	$48\mu s$	$2\mu s$
-15°	105°	$85\mu s$	$53\mu s$	$2\mu s$	115°	$117\mu s$	$50\mu s$	$2\mu s$
0°	105°	$84\mu s$	$54\mu s$	$2\mu s$	115°	$119\mu s$	$52\mu s$	$3\mu s$
15°	105°	$92\mu s$	$57\mu s$	$2\mu s$	115°	$108\mu s$	$52\mu s$	$2\mu s$
30°	60°	$61\mu s$	$52\mu s$	$1\mu s$	115°	$85\mu s$	$50\mu s$	$1\mu s$
45°	90°	$61\mu s$	$56\mu s$	$0\mu s$	160°	$59\mu s$	$47\mu s$	$1\mu s$
60°	90°	$65\mu s$	$58\mu s$	$1\mu s$	155°	$51\mu s$	$43\mu s$	$1\mu s$

Table 2: Peak ITD asymmetry values and curve summarization (asymmetry curve mean and standard error) across examined databases, for different elevations. Values are approximated to the nearest microsecond.

as well as their different measurement techniques, prevented a pooling approach. For this paper, the FIU database [10] was removed from the analysis due to the poor azimuthal resolution of data (caused by errors in the stored repository). The MARL database [11], which consisted of 32 sets of repeated measurements for four subjects, was also removed since the different HRIR measurement techniques used within the set were deemed to create excessive ITD variability.

The intent of this paper is to examine whether the asymmetry curve at different elevations follows a similar pattern to that observed in the horizontal plane (Figure 1). ITDs and ITD asymmetry were calculated using the same methods used for the case of the horizontal plane (see Section 3). The elevations considered ranged to what was available for every database within the range $\phi \in [-45°, +60°]$ in 15° increments (for ITD extraction, CIPIC elevations were approximated to the nearest elevation to the increment). For clarity, only elevations of $\phi = \{-45°; -30°; 0°; +30°; +45°\}$ are included in the graphs, while full-range asymmetry peak values are included in table 2.

Figures 2 and 3 show the average ITD asymmetry between hemispheres in microseconds for different elevation angles in the LISTEN and CIPIC databases respectively. The main region of interest is the one previously identified as "sensitivity region" highlighted in Figure 1, which spans between $\theta \in [\pm 90°, \pm 130°]$. The general observation is that of a "flattening" pattern of the curves within this region. While this process seems to happen more or less linearly for changing elevations $|\phi|$ in CIPIC (Figure 3), it does not occur as clearly in the LISTEN dataset (Figure 2), although the curves also follow a flattening trend.

Table 2 provides details of the azimuth location at which the peak asymmetry occurs for each elevation. The table includes the mean asymmetry and standard error of the mean for each curve, used to characterize the curve in relation to the peak value. It is worth noticing that for both datasets there is little change in proximity of the zero elevation curve ($\phi = +15°$ and $\phi = -15°$). As shown in the graphs, elevations moving away from the horizontal plane ($\phi = 0$) cause a decreasing trend in the magnitude of the peak asymmetry with respect to the average curve value, further indicating how the region of sensitivity levels to the rest of the curve.

5 Discussion

The produced plots for the two datasets generally agree with each other, but their interpretation must be accompanied by a consideration of Table 2. Increasing or decreasing the elevation angle would reduce the peak levels of the asymmetry in the sensitivity region for azimuth between $\pm 90°$ and $\pm 130°$. By visual inspection, the general trend observed in Figure 2 (LISTEN) is a flattening of the sensitivity region as the elevation moves away from the horizontal plane. The $\phi = -45°$ case presents a peak at $\theta = 75°$ which is unusual. However, in this set, the general "flattening" of the region seems to happen independently of the direction of elevation change. This is not exactly the case for Figure 3 (CIPIC) where the direction of elevation change seems to influence the rate of change of the flattening process. The most interesting aspect of this graph is how the region of sensitivity is still recognizable across elevations but gradually levels with the rest of the curve as the distance from the horizontal plane linearly increases. In other words, the ITD asymmetry peaks decrease linearly as the elevation angle moves away from $\phi = 0$.

According to Table 2, where the values of the curves' peaks are examined with reference to the curves' mean, it is possible to notice how increasing or decreasing the elevation "flattens" the peak region and the overall curve towards a "floor" value. This trend is observed to happen almost linearly for the CIPIC dataset but not as much for LISTEN. One could suggest that when elevating a sound source position, its interference with the outer ear is reduced. For decreasing elevations, the situation is more complex and the sensitivity region does not decrease as fast, which could be explained by the fact that the listeners' torso might add a new layer of interference between the sound source direct path and the ear canal entrance.

More reliable population asymmetry curves could be obtained if the data had a uniform resolution format that permits pooling the ITDs into a single analysis. The non-uniformity of the datasets' collection methods and angle resolutions make it hard to draw confident conclusions about the statistical presence of ITD asymmetry across the listener population. Due to the low temporal resolution, small ITD errors find various possible causes along the stages of the measurement/processing chain. Errors on the order of one or two samples might be the cause of the general "floor" asymmetry level approximately around $50\mu s$ (each sample corresponds to a time interval of $22.6\mu s$). A more rigid standardization of measurement techniques and repositories as proposed in [11], and a faster sampling rate, could improve the reliability of statistical exploratory studies, such as this one.

The question of whether the measured ITD asymmetries are perceptually relevant for the average listener is not currently addressed by this paper. However, literature indicates that the asymmetries might indeed be significantly noticeable. A study by Klumpp and Eady [12] found the ITD *Just Noticeable Difference* to be frequency dependent between $11\mu s$ to $95\mu s$ for pure tones, while a random noise signal would have a JND of $9\mu s$. In Figure 1 we can see that for the horizontal plane, the peak asymmetries are all above the range defined by Klumpp. This means that for the region of sensitivity the asymmetry would be above the noticeability threshold, if not compensated for. The values observed in Figures 2 and 3 are also above the JND threshold for most frequency bands. By combining this information with the patterns seen in the plots, the perceptual noticeability of the ITD asymmetry should decrease as the elevation moves away from zero.

6 Conclusions

This study extended the analysis of ITD asymmetry across publicly available datasets of individual HRIR measurements to the elevation dimension. Two datasets were chosen for their widespread use and precise alignment technique. ITD asymmetry was calculated as the average difference, across a set of subjects, between the ITD magnitude at azimuth θ and its corresponding azimuth at the opposite hemisphere (Equation 1). Although some of the asymmetries are likely generated by measurement error and exacerbated by a poor sampling rate, the general trend is that elevation affects the severity of the asymmetry.

A previously identified sensitivity region between $\theta = \pm 90°$ to $\pm 130°$ was found to be affected by a change in elevation, and tends towards a flatter curve which stabilizes around a "floor" asymmetry value around $50\mu s$. The combined information of Figures 2 and 3 together with Table 2, suggest that the sensitivity region is flattened by a change in the elevation angle independently from the direction of change. However, the two sets differ in the way in which this change happens. The CIPIC dataset (Figure 3) shows that the peak and mean asymmetry decreases linearly as $|\phi| > 0°$ increases. For the LISTEN set (Figure 2) the change is not as linear, although it still points towards a flattened curve. In summary, increasing or decreasing the elevation angle will reduce the impact of morphological asymmetries on ITD measurements, possibly in a linear fashion. If we assume that part of the recorded asymmetry is due to measurement or quantization error, it seems that ITDs become more symmetrical as the elevation increases, especially above $\phi = 0°$.

The peak ITD asymmetry values for all elevations are roughly near the same sensitivity region found in [7]. All the peak values are above the minimum JND levels found by [12], but below the maximum range, suggesting that the lower frequency bands would not be affected by the recorded asymmetries. However, results are not yet validated by a perceptual study which would examine the actual impact of asymmetries in HRIRs for precise localization tasks. Such a test should improve our understanding of the importance of asymmetric spatial cues.

6.1 Future Work

It is not yet clear which particular anthropometric parameter most significantly affects symmetry. A correlation study could possibly highlight which feature should be regarded as the most influential. Table 1 suggests that the "Pinna Height" is a candidate for this role since it is the feature with the most variability within the CIPIC database. However, public databases are not usually accompanied by detailed anthropometric parameters for the measured subjects, making such a study hard to perform.

Results have not yet been validated by a perceptual test which would further assess the noticeability of the ITD asymmetries and the correlation between measured and perceived asymmetry. Neural compensation caused by internal-delay has been hypothesized to be a factor in spatial perception [13]. If a correlation were to be found, then a correction parameter calculated from morphological asymmetries could be implemented within an HRIR individualization model.

Current modeling techniques already make an effort to integrate anthropometric parameters for simulating individual HRIRs, for example through photographic techniques [5]. A recent extension of the spherical head model was developed by [14] to take account of the elevation dependance of the model using simple head measurements, and therefore better approximate front/back ITD asymmetry. However, lateral asymmetry across hemispheres at different elevations is not addressed. To confront this issue it is required to more carefully observe head/pinnae parameters for both head hemispheres and consider any offsets from the axis of symmetry. It is currently not trivial to do so, but it is possible to imagine a future where comprehensive photographic feature-collection techniques will be able to extract these detailed features using commercial devices. Furthermore, to identify which morphological feature could more strongly affect the ITD symmetry could considerably narrow down the complexity of such proposed individualization technique.

References

[1] Morimoto, M. and Ando, Y., "On the simulation of sound localization." *Journal of the Acoustical Society of Japan (E)*, 1(3), pp. 167–174, 1980.

[2] Wightman, F. L. and Kistler, D. J., "The dominant role of low-frequency interaural time differences in sound localization," *The Journal of the Acoustical Society of America*, 91(3), pp. 1648–1661, 1992.

[3] Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C., "The cipic hrtf database," in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, pp. 99–102, IEEE, 2001.

[4] Dellepiane, M., Pietroni, N., Tsingos, N., Asselot, M., and Scopigno, R., "Reconstructing head models from photographs for individualized 3D-audio processing," in *Computer Graphics Forum*, volume 27, pp. 1719–1727, Wiley Online Library, 2008.

[5] Juras, J., Miller, C., and Rogińska, A., "Modeling ITDs Based On Photographic Head Information," in *Audio Engineering Society Convention 139*, 2015.

[6] Zhong, X.-L., Zhang, F.-c., and Xie, B.-S., "On the spatial symmetry of head-related transfer functions," *Applied Acoustics*, 74(6), pp. 856–864, 2013.

[7] Genovese, A. F., Juras, J., Miller, C., and Rogińska, A., "Investigation of ITD symmetry

across existing databases of personalized HRTFs," in *SUBMITTED TO ICAD convention*, 2016.

[8] Boehnke, S. E. and Phillips, D. P., "Azimuthal tuning of human perceptual channels for sound location," *The Journal of the Acoustical Society of America*, 106(4), pp. 1948–1955, 1999.

[9] Warufsel, O., "LISTEN HRTF Database, IRCAM," 2002.

[10] Gupta, N., Barreto, A., Joshi, M., and Agudelo, J. C., "HRTF database at FIU DSP lab," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 169–172, IEEE, 2010.

[11] Andreopoulou, A. and Rogińska, A., "Towards the Creation of a Standardized HRTF Repository," in *Audio Engineering Society Convention 131*, 2011.

[12] Klumpp, R. and Eady, H., "Some measurements of interaural time difference thresholds," *The Journal of the Acoustical Society of America*, 28(5), pp. 859–860, 1956.

[13] Joris, P. and Yin, T. C., "A matter of time: internal delays in binaural processing," *Trends in neurosciences*, 30(2), pp. 70–78, 2007.

[14] Sridhar, R. and Choueiri, E., "Capturing the Elevation Dependence of Interaural Time Difference with an Extension of the Spherical-Head Model," in *Audio Engineering Society Convention 139*, 2015.