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

In binaural simulations, *Head-Related Impulse Responses* are used to recreate a 3D auditory display through headphones. Public repositories of individually measured HRIRs are widely used in industry and research. However, head-related anthropometric asymmetries, among measured subjects, are a likely cause of measured asymmetries in *Interaural Time Delay* cues (ITDs), which may lead to imprecise sound localization. As part of a larger study on HRIR personalization, this paper expands, to the elevation dimension, the investigation of ITD asymmetry in public databases of measured HRIRs. In a previous exploratory study, concerning the horizontal plane only, a *region of sensitivity*, where the ITD asymmetry was observed to be significantly more prominent, was identified in datasets of individually measured HRIRs approximately between the azimuth range of $\theta = \pm 90°$ to $\pm 130°$. For this paper. two publicly available databases of individual HRIRs were selected and analyzed in search of an elevation effect on ITD symmetry. Results found that an increase or decrease in elevation angle ϕ, away from the horizontal plane, affects the asymmetry curve by reducing the gap between average and peak ITD asymmetry values within the mentioned region in a roughly linear trend. This finding points to the fact that, within the examined datasets, the statistical presence of ITD asymmetries is gradually less severe, although still present, as the elevation angle moves away from the horizontal plane.

1 Introduction

In 3D binaural applications, 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, which is then typically played back over a headphone stereo output. 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 are the primary cues for low frequencies [1].

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 head-related morphological parameters, i.e. the shape of the outer ears, match those of the used dummy head. Furthermore, human heads and outer ears are seldom symmetric [2]. Asymmetry in pinnae and head shape might imply, for some listeners, a significant asymmetry in the physical spatial cues received at the ears, in particular, ITDs. For these reasons, individualized HRIRs, measured directly on the listener, are considered to be desirable for a user when experiencing binaural sound.

However, since it is not feasible to measure HRIRs for the consumer market, current modeling techniques already make an effort to integrate anthropometric parameters when synthesizing or adapting individual HRIRs, for example through photographic techniques using commercial devices [3] [4]. A recent method extends a commonly used spherical head model to take account of the impact of the elevation angle on the head-model using simple anthopometric measurements [5], therefore better approximating front/back ITD asymmetry. However, lateral asymmetry across hemispheres at different elevations is not addressed. In fact, HRIR modeling techniques often assume a symmetrical head, meaning that the spatial cues recreated for one of the lateral 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 binaural auditory displays.

1.1 Motivation

A previous paper by the authors [7] showed the existence of relevant ITD asymmetry between the left and right hemispheres on the horizontal plane ($\phi = 0°$). In the four examined databases, and in the pooled sample, a common azimuthal region of increased asymmetry was observed and suggested to be approximately between $\theta \in [90°, 130°]$. The analysis revealed a statistically significant difference between the peak asymmetry within the region and the average asymmetry across all other angle pairs. Furthermore, the peak asymmetry values were found to be much higher than the established (*Just Noticeable Difference*) JND values for localization discrimination, at all frequency bands [8].

A possible explanation, due to be further explored, is that human head/ear asymmetries, rather than subject misalignment, would likely be the cause of the aforementioned region of augmented ITD asymmetry. The CIPIC database includes complementary anthropometric data for 37 out of the 45 measured subjects. A preliminary analysis on the symmetry of this anthropometric data was performed by averaging, across subjects, the difference between left and right features. The parameter that presented the most variability was *"Pinna Height"*, which showed 0.3207 cm average difference between left and right ears. It is not yet clear how severe is the impact of anthropometric asymmetries on ITDs asymmetry, and whether they have a directional effect, but their presence supports the hypothesis that mirrored ITDs are asymmetric regardless of measurement error. The implication would be that current and future HRIR modeling techniques need to consider the implementation of asymmetrical ITD cues if the aim is to achieve accurate sound localization ability. This issue is therefore relevant for precision-dependent applications of binaural audio that require a high-degree of accuracy. Some example applications are data sonification displays or navigation simulations.

The motivation behind the present analytical study is primarily to further explore the statistical presence of asymmetric ITDs in public HRIR repositories (commonly used for modeling and research) by extending the analysis to the elevation dimension, and secondarily to initiate a discussion on the role of morphological asymmetries in ITD asymmetry. Two publicly available databases (also examined in [7]) of individual HRIR measurements are analyzed and inspected for elevation-dependent asymmetry between locations mirrored across the left and right hemispheres.

2 Databases

The data for this analysis was collected from two publicly available databases of individual HRIRs. The two selected databases, LISTEN [9] by IRCAM (50 subjects) and CIPIC [2] by UC Davies (45 subjects), differed in angle resolution and measurement technique. Due to these differences, it was decided to analyze the two databases also independently, in addition to a merged interpolated repository. Both sets were recorded at a sampling rate of $44.1kHz$ which corresponds to a temporal resolution of $22.6\mu s$ and represents the minimum detectable ITD asymmetry value above zero.

2.1 LISTEN database

The HRIRs in this database were recorded by IRCAM and AKG for their collaborative LISTEN project [9]. HRIRs were measured for variable azimuth resolutions at different elevations (typically with 15° increments) with elevation $\phi \in [-45°, +90°]$ (also, 15° increments) [1]. The measurements were conducted on 50 individual subjects for a total of 187 locations per subject at $44.1kHz$. The HRIRs used for this analysis were used in their *compensated* form, hence equalized and time aligned for zero frontal ITD by the author of the set.

The measurement technique made use of a very precise crane structure connected to a single loudspeaker able to move across a rig, and a software controlled rotating chair provided with a headrest. This methodology ensures a satisfactory degree of reliability in the subject's azimuth alignment precision and is generally less prone to errors caused by human misalignment.

2.2 CIPIC database

UC Davis measured 45 subjects at 1250 spherical locations for 25 non-uniformly-spaced azimuth locations where $\theta \in [-80°, +80°]$ and 50 elevations $\phi \in [-45°, +230.625°]$ (step increments of 5.625°) [2]. Golay codes at $44.1kHz$ 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 dataset is provided with complementary anthropometric data for some of the measured subject. The anthropometric data consists of a detailed collection of measurements, for both hemispheres, of head and pinnae parameters such as head size parameters, pinnae features dimensions and ear rotation angles.

3 Previous Findings

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 measurements. In each of the four used dataset, the average asymmetry $\bar{S}(\theta)$ for horizontal plane elevation ($\phi = 0$) 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 azimuth angle where a counterpart existed, thus excluding $\theta = 0°$ and $\theta = 180°$).

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

During the pre-analysis stage, the databases were interpolated to a common azimuthal resolution of 5° steps and removed of strong outliers. Figure 1 shows the average ITD asymmetry curve across the horizontal plane for both the two databases taken individually, and a pooled repository. The resulting graphs show an interesting common *region of sensitivity*, roughly between 90° and 130° azimuth, where the average ITD asymmetry seems more accentuated. This hypothesis was confirmed by t-tests performed between the peak asymmetry and average asymmetry, which showed a statistically significant difference between the two, at a 99% confidence level (for both the individual sets and the pooled repository). Although a perceptual validation study on the impact of these findings has yet to be performed, the observed severity of the asymmetries was found to be larger than established JND values at all frequency bands [8]. In fact, the peak values of ITD asymmetry ranged between $80 - 94\mu s$ while the JND for a 90Hz pure tone is set to $75\mu s$ and decreases as the frequency increases (band-limited random noise presents a JND value of $9\mu s$). By making these comparisons, it is possible to infer that asymmetric ITDs inside these datasets can be perceptually noticeable, making the issue relevant to the use of these HRIRs in research and binaural simulation.

The main conclusions established the fact that commonly used public repositories do not contain symmetric ITDs, and there is a presence of a significant asymmetric region which magnitude is above JND values for spatial perceptual discrimination. These findings are currently relevant to parties interested in public HRIR sets, and might find further confirmation in future dataset collections. Furthermore, although it is expected that a number of these asymmetries are the fruit of measurement errors and human misalignments, the presence of a significantly more asymmetric region

[1] For this paper only elevations up to +45° were considered

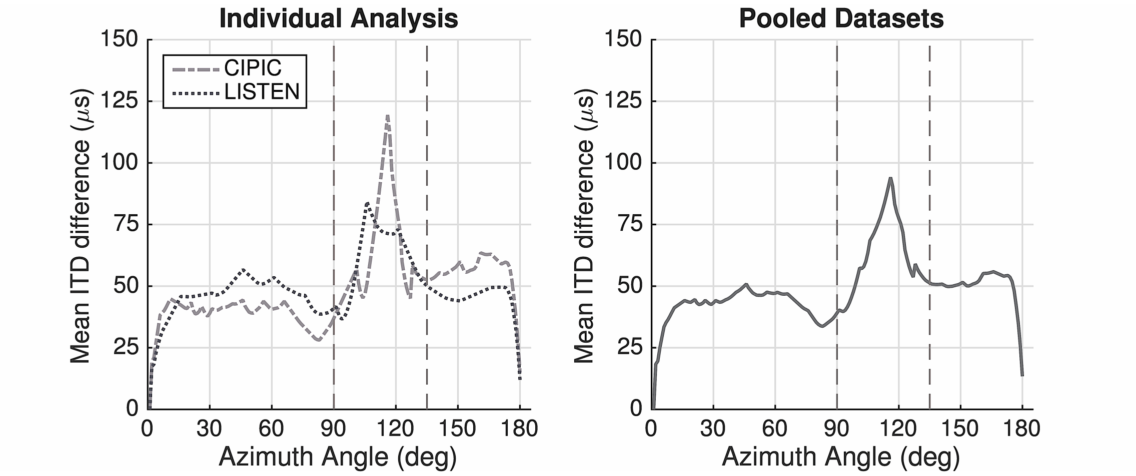

Fig. 1: Average ITD asymmetry across the horizontal plane. The first plot illustrates the differences between the LISTEN and CIPIC sets. The second plot shows the results for the pooled merged datasets. An identified region of sensitivity spans between 90° and 130° (figure from [7])

results raise the hypothesis of a directional-dependent influence of anthropometric asymmetries, the alternative being an unlikely systematic directional error, coincident in both datasets.

4 Analysis and Results

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 the symmetry values were calculated using the same methods and equations used for the horizontal plane investigative study (see Section 3). Differently from the previous study [7] which used four sets, only two HRIRs datasets were picked for this study. CIPIC and LISTEN were chosen due to their reliable measurement technique and the availability of a comprehensive sphere of locations. Both sets were interpolated to every 5° azimuth, to account for the different resolutions, and merged into a single set, for each available common elevation, which ranged from $\phi \in [-45°, +45°]$ with step increments of 15°.

Figure 2 shows the average ITD asymmetry between hemispheres in microseconds, for different elevation angles. For convenience, results were plotted for all elevations above zero in the first graph, and below zero in the second. The dashed vertical lines represent the previously identified *sensitivity region*. Out of preliminary visual inspection, the general observation is a "flattening" pattern of the curves as the elevation drives away from the horizontal plane. By looking at the lower graphs, it seems that the flattening trend happens more or less linearly with the change in elevation, with the $\phi = -45°$ and $\phi = +45°$ curves being the most *flat*. The location of the peak asymmetry value remains unvaried throughout the elevation span.

Table 1 provides the details of the peak asymmetry values, peak location, and mean curve asymmetry for each studied elevation. The table also includes the values for both sets taken independently. 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 ($\phi = 0$) generally 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. Some other observations that can be drawn from the table, are the facts that the CIPIC dataset possesses the highest asymmetry peaks for most elevations, while the LISTEN dataset has a slightly higher average asymmetry.

By means of *t-tests*, it was possible to demonstrate that for some elevations, closer to zero, the ITD asymmetries within the *sensitivity region* are statistically larger than the average, while for the elevations farther than zero, they are not. Two-tailed t-tests were conducted at 99% and 95% confidence level between

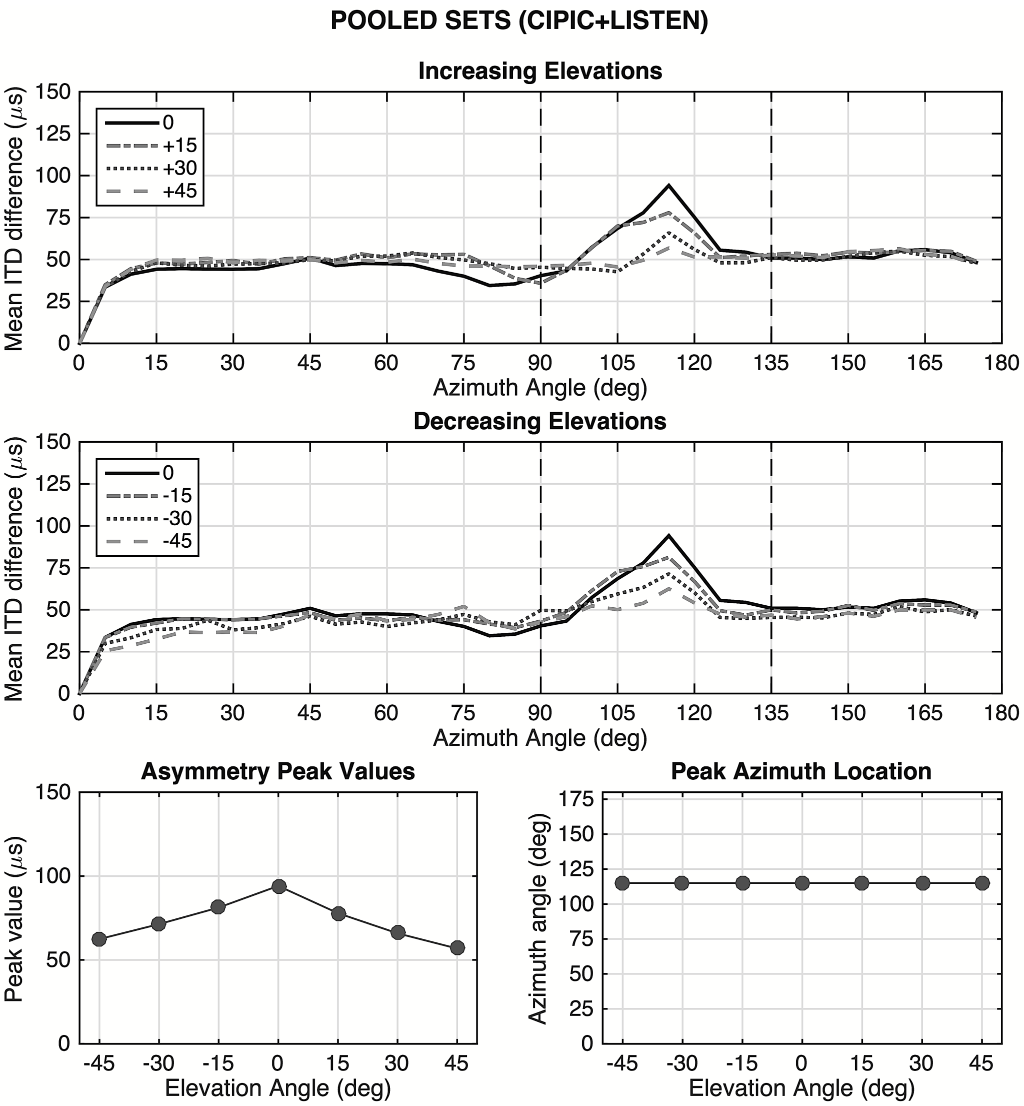

Fig. 2: Elevation curves for the pooled repository (CIPIC+LISTEN). The lower graphs show the change in peak asymmetry and position across different elevations

Database	**POOL**			**LIST.**			**CIPIC**		
El. (ϕ)	Peak	Location	Mean	Peak	Location	Mean	Peak	Location	Mean
+45°	57μs	115°	49μs	61μs	90°	54μs	59μs	160°	44μs
+30°	66μs	115°	49μs	61μs	60°	52μs	**85μs**	115°	48μs
+15°	**79μs**	115°	52μs	**92μs**	105°	54μs	**100μs**	115°	50μs
0°	**94μs**	115°	51μs	**84μs**	105°	51μs	**119μs**	115°	51μs
-15°	**81μs**	115°	50μs	**85μs**	105°	50μs	**109μs**	115°	49μs
-30°	**71μs**	115°	46μs	61μs	75°	45μs	**100μs**	115°	47μs
-45°	62μs	115°	45μs	73μs	75°	44μs	80μs	115°	45μs

Table 1: Peak ITD asymmetry values, peak azimuth locations, and asymmetry curve means across examined databases, for different elevations. Values are approximated to the nearest microsecond. Peak values in boldface showed a statistical significant difference from the curve mean value at 99% confidence level *($p < 0.01$)*.

Database	**POOL**			**LIST.**			**CIPIC**		
El. (ϕ)	*p-value*	< 0.01	< 0.05	*p-value*	< 0.01	< 0.05	*p-value*	< 0.01	< 0.05
+45°	0.0542	✗	✗	0.1476	✗	✗	0.0484	✗	✓
+30°	0.01320	✗	✓	0.2759	✗	✗	0.0042	✓	✓
+15°	0.002	✓	✓	6.44e-06	✓	✓	0.0016	✓	✓
0°	9.36e-06	✓	✓	1.79e-04	✓	✓	2.8e-04	✓	✓
-15°	5.82e-04	✓	✓	9.22e-04	✓	✓	8e-04	✓	✓
-30°	0.0033	✓	✓	0.034	✗	✓	0.0019	✓	✓
-45°	0.0149	✗	✓	0.2830	✗	✗	0.0133	✗	✓

Table 2: Details of significance analysis for 99% and 95% confidence level for POOL *(df = 94, N = 95)*, LISTEN *(df = 49, N = 50)* and CIPIC *(df = 44, N = 45)*.

the peak asymmetry within the region $\theta \in [90°, 130°]$ and the average asymmetry for the rest of the curve. For all three sets, POOL *(df = 94, N = 95)*, LISTEN *(df = 49, N = 50)* and CIPIC *(df = 44, N = 45)*, the test showed significance *(p < 0.01)* for the elevations close to zero ($\phi = 0°, -15°, +15°$). The pooled repository and CIPIC showed significance also for elevation $\phi = -30°$, while only CIPIC showed it for $\phi = +30°$ (details in Table 2). The trend is the same for all three cases, the region becomes less prominent as the elevation moves farther apart from the horizontal plane.

5 Discussion

The most interesting aspect of this study is how the peak region gradually levels with the rest of the curve as the distance from the horizontal plane increases or decreases. In the pooled repository, the ITD asymmetry peak decreases almost in a linear fashion as the elevation angle moves away from $\phi = 0$. The observed process does not happen as linearly for the LISTEN set when taken individually, although it is, generally, still in agreement with the trend. The effect of the elevation change is validated by the t-test results, which showed a common significant difference, between the peak region and the average asymmetry, only for the elevations close to zero, in both directions. On Table 1, the values of the curves' peaks should be examined with reference to the curves' means. It is then possible to validate how increasing or decreasing the elevation, "flattens" the peak region and the overall curve towards a general *floor* value, which could be interpreted as a base measurement error, or misalignment tolerance. In fact, with the temporal resolution given by the databases' sample rate, small ITD errors on the order of one or two samples might be the cause of the floor asymmetry level, approximately around $50\mu s$ (at $44.1kHz$ each sample corresponds to a time interval of $22.6\mu s$).

Given the reliability and professionality of the measurement procedures used for these databases, it is hard to declare that the region of increased asymmetry is caused by a coincident systematic measurement error. Assuming directional equiprobability of error and human misalignment across azimuths and elevations, the significant region of accentuated asymmetry could be attributed to head-related morphological asymmetries having a directional-dependent influence on asymmetric ITDs in measured HRIRs. In the elevation case, it is possible to hypothesize that this influence softens as the interaction between the angle of incidence of the sound source and the outer ear is reduced, causing the observed flattening trend.

Literature indicates that the listed asymmetry curves are above the indicated JND values for perceptual noticeability at most bands, regardless of elevations, since the mean asymmetry is not much affected. The ITD JND values indicated by Klumpp and Eady [8] for spatial discrimination are between $11\mu s$ to $95\mu s$ for pure tones, and $9\mu s$ for a random noise signal. It is clear that although greater elevations might reduce the sensitive region, the floor asymmetry is still within the perceptually noticeable range. Hence, the ITDs in the examined databases have the potential to be perceptually inaccurate for most listeners, except, perhaps, the lowest frequency bands. In 3D binaural, the importance of the directional mismatch between hemispheres would be dependent on the frequency content as well as the purpose of the application. Binaural simulations that look for a very high degree of localization accuracy would have to compensate for this issue with a correction parameter, although other factors such as training and asymmetric perception [10] might come into play. At least for the examined databases, the change in elevation does not eliminate measured ITD asymmetry, but reduces the impact of the most severe regions.

6 Conclusions

This study extended to the elevation dimension the analysis of ITD asymmetry across publicly available datasets of individual HRIR measurements. Two public datasets were chosen for their widespread use in research and relatively reliable alignment technique. The results highlighted a "flattening" trend in the ITD asymmetry curves as the elevation moves away from the horizontal plane, independently from the direction of change. Following this trend, the *sensitivity region* previously identified [7] as between the azimuth range of $\pm 90° < \theta < \pm 130°$ was demonstrated to lose prominence and significance in a roughly linear fashion. Despite the reduced levels of the peak asymmetries in the region, all curves present *floor* levels of asymmetry which average around $50\mu s$, potentially caused by measurement and quantization errors in addition to morphological asymmetries in measured subjects.

The average ITD asymmetry values for all elevations are similar, and above the JND levels for spatial discrimination except for lowest frequency tones [8], suggesting a frequency-dependent perceptual impact. However, for the horizontal plane, and the elevation planes close to zero, the peak asymmetries surpass JND levels at all frequency bands. These findings are relevant to the fields of HRIR modeling and HRIR personalization for accurate binaural simulations.

6.1 Further Work

The extent of the impact of head-related morphological asymmetries in measured HRIRs remains to be further explored as more data is added to the available repositories. The current exploratory study would benefit from the addition of new datasets from both industrial and academic origin. A more rigid standardization of measurement techniques and repositories (as proposed in [11]), and a faster sampling rate, leading to better temporal resolution, could improve the reliability of statistical exploratory studies, such as this one.

As this study stands as part of a bigger research effort on individualization techniques, the next step would be to investigate on which particular anthropometric parameter most significantly affects ITD symmetry. 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. A correlation study could possibly highlight which feature should be regarded as the most influential. It has been suggested that "Pinna Height" is a candidate for this role since it is the feature with the most variability within the CIPIC database. If correlations were to be confirmed, then a correction parameter calculated from morphological asymmetries could be implemented within an HRIR individualization model. However, public databases are not usually accompanied by detailed anthropometric parameters for the measured subjects, making such a study not trivial to perform.

As of now, results are not yet validated by a perceptual study, which would further 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 and further assess the noticeability of ITD asymmetries.

References

[1] 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.

[2] 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.

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

[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] 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.

[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 in measured HRIRs," in *22nd International Conference on Auditory Display, Canberra*, 2016.

[8] 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.

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

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

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