Audio Engineering Society

Convention Paper 9754

Presented at the 142nd Convention
2017 May 20–23 Berlin, Germany

The evaluation of the effect of sound directionality in horizontal plane on human auditory distance perception in a large reverberant room

Tahereh Afghah[1], Andrew Allen[2], Peter Otto[1] and Aravindan Joseph Benjamin[3]

[1] *University of California San Diego, San Diego, California, USA*

[2] *Google Inc., Mountain View, California, USA*

[3] *Technical University of Ilmenau, Ilmenau, Thuringia, Gemany*

Correspondence should be addressed to Tahereh Afghah (tafghah@ucsd.edu)

ABSTRACT

An evaluation of sound localization effect on the auditory distance estimation in a user study is presented. Binaural Room Impulse Responses of sixty positions were recorded in a reverberant space using a dummy head. The recordings were evaluated by the users in a headphone-based listening test to analyze the listeners' ability to perceive the distance with and without prior knowledge of direction of origin. When known, the distance estimation accuracy in left and right sides of the head in near field (2m,4m) was improved and at some angles saw a significant improvement. However, known direction did not assist the users in determining the larger distance levels (6m,8m,10m). No improvements were seen in the front and back sides for all directions.

1 Introduction

The human auditory system functions as an "early warning system" to recognize threats from the spatial surrounding environment. By recognizing angle and distance cues, the auditory system enables spatial awareness of these events [1]. With the advent of Virtual Reality technology, it becomes possible to simulate real-world sensations by synthesizing virtual auditory stimuli using these cues. This enables us to perform tasks such as training visually impaired persons to distinguish incoming threats or to enhance content for Virtual Reality entertainment applications. Many studies have shown that head-related transfer functions (HRTFs) produce salient timing and level differences that are important for auditory localization and primarily assist in perceiving the direction of incoming sound [2] [3] [4] [5]. Distance cues are provided by relating the sound arriving directly at the listener from the source without any reflections with those that arrive after reflecting off of surfaces inside the listening environment. The response of a room plays a large role in determining the angle and distance of incoming sound. In an enclosed room, reflections off walls can diffuse the sound and blur the direction of origin [6] [7] [8]. In this study, the effect of direction of incidence on auditory distance perception in a large reverberant room was evaluated. Sixty different positions were specified and the sound coming from these positions was recorded. Then a psychoacoustic listening test was performed for the assessment on this influence.

2 Auditory distance perception

There are different static distance perception cues that influence the accuracy of the distance estimation, depending on the situation of the propagation environment. Many studies have focused on the main cues including sound intensity [7] [9], reverberation [6] [10], Interaural Time and Level Differences (ITD and ILD) and the frequency spectrum of the sound, but there are other cues that can also have an effect on auditory distance perception [11] [12] [13]. Additionally, in case of having either a moving sound source, moving receiver (microphone or ears), or both during the recording or play back of the sound, dynamic distance perception cues, including motion parallax and acoustic tau would also have an influence on the results [9] [14].

The familiarity of the target sound, existence of background noise, the temperature and humidity of the sound propagation environment, vision cues available to the listeners, apparent sound source width and the direction of the received sound can also be considered as valid cues on top of the main cues. More investigation is needed to determine when it is appropriate to consider these cues as main distance perception cues [9] [15] [16] [17].

3 Method

In this study, a listening test was designed to evaluate the direction of incidence effect on auditory distance perception. In applications such as AR or VR, the listener is not aware of the room features of the recording. Therefore, in this research, the recording and play back process were done in two separate rooms to evaluate this effect without listener's potential learning due to the prior knowledge of the recording situation. This is beneficial to ensure the visual cue including the dimensions of the recording space did not affect the localization accuracy. To maximize the sound source diffusion, a large reverberant space with a long decay time was chosen for the recording process. The research was conducted through three main steps.

3.1 Recording

The first step was aimed at establishing a loudspeaker layout that can be used efficiently to evaluate localization. To do so, a dummy head[1] was positioned at the center of a large room[2] with the following dimensions : 36 x 27 x 9 m. All the recordings for this experiment were done using an in-ear microphone pair[3] which was positioned inside the dummy head. Although using individualized BRIR instead of recording with a dummy head could yield more precise localization, it was not practical to repeat the recording process for each subject as in the AR and VR application's design. Furthermore, as this procedure was so time consuming, individual recordings could cause listener fatigue.

Recordings were made from sixty separate source locations using one loudspeaker[4] that was moved to each position successively. Five concentric circles with radius 2, 4, 6, 8 and 10 m, centered around the dummy head. In each circle, 12 positions 30 degrees from one another were defined (Fig. 1).

The position of the speaker was carefully calibrated using laser measurements[5]. The height of the speaker was adjusted in order to be horizontally aligned with the dummy head. The target sound consisted of two seconds of wide-band white noise [20 - 20000 Hz] followed by five seconds of silence in order to accommodate the estimated "3.7 s" RT-60 of the room. Fig. 2 represents the temporal shape of the target signal. The burst was played through the speaker and the corresponding response was recorded for each location separately. Fig. 3 shows the frequency response of the recordings at 0 degrees 2, 4, 6, 8 and 10 m away from the dummy head.

3.2. GUI design

The second step involved designing a listening test to record the response of the test subjects and gather the data. This test consisted of a graphical user interface (GUI) designed in MATLAB. The GUI contained training and question sections.

---

[1] Bruel & Kjaer, 4100D
[2] Location: The University of California San Diego - Price center - Ballroom AB
[3] 4189-A-002 free-field microphone
[4] Genelec A 8030
[5] Johnson Level Tool 40-0921

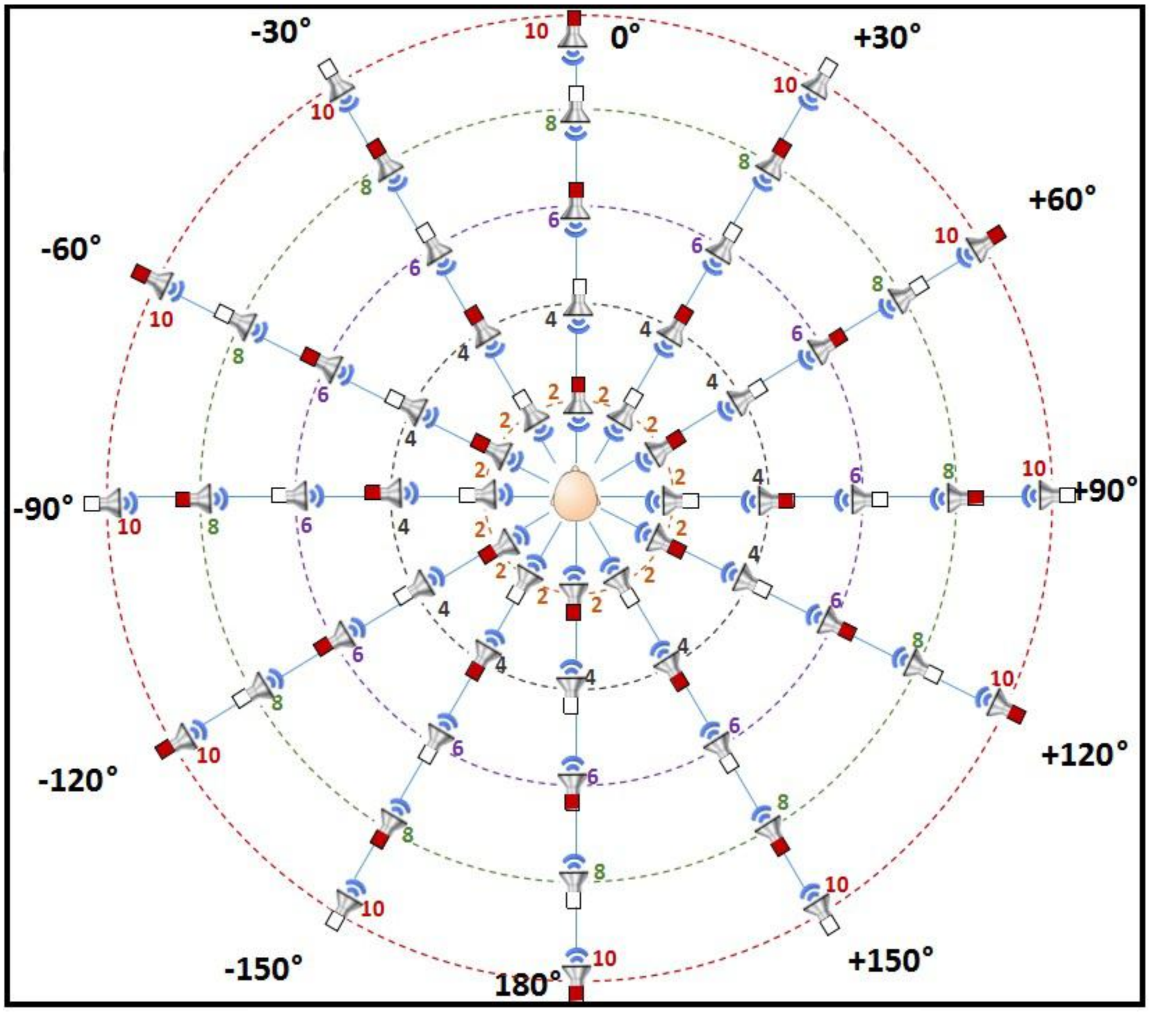

Figure 1. The main grid of the recording set up

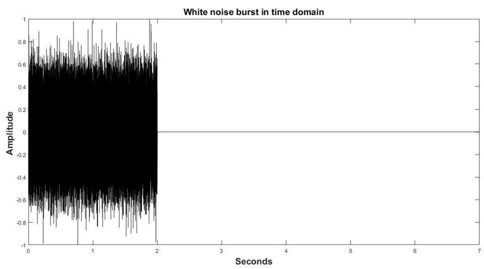

Figure 2. The white noise burst representation

To ensure all the subjects were trained in the same way, they were not able to control the order of play back or the number of repetitions during the entire training section. Three main phases were considered in designing the training section. In each phase, the sound coming from any position was played only once.

3.2.1 Training phase 1

The first phase taught the concept of the "direction" to the listeners. A diagram showing all the possible positions of the speakers was displayed on the screen. This phase was started by playing the target sound coming from the speaker positioned at 2 m and 0 degrees from the dummy head while showing only the corresponding speaker icon on the screen. Next, the following speaker at 2 m and 30 degrees was shown and the corresponding recorded sound was played. This process was repeated for all

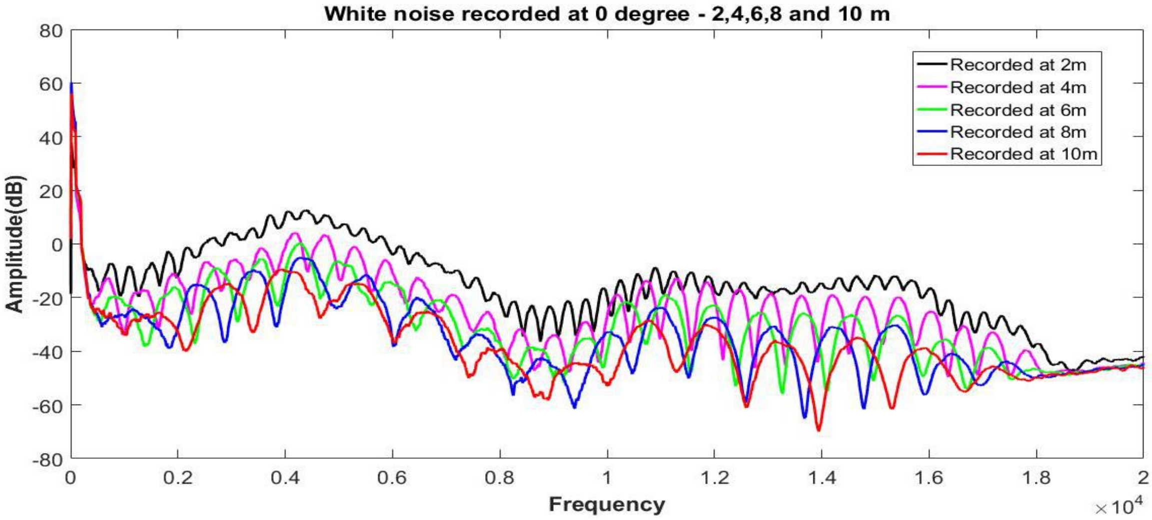

Figure 3. The frequency responses at 0 degrees

12 speakers 2m away. This entire process was then done with the speakers arranged on the other circles. This was to emphasize how the sound was changing when coming from the same distance, but different directions.

3.2.2 Training phase 2

The second phase taught the concept of the "distance". It began by playing the sound coming from the speaker positioned at 2 m and 0 degrees from the dummy head. Next, the sound coming from the source positioned at 4 m and 0 degrees was played. The same distance order was considered for all the 12 directions. This was to emphasize how the sound was changing when coming from the same direction, but different distance levels.

3.2.3 Training phase 3

The third phase of the training section was a sample test including the localization of twelve out of sixty positions which were the same for all test subjects. For each position, all sixty speaker icons were shown on the screen while the target sound was played from one position. The subject was asked to listen to the sound and choose the source position by clicking on the corresponding speaker shown on the screen. The GUI recorded the chosen answer and displayed the correct answer afterwards. The outcome of this phase was not included in the final results as it was performed merely to familiarize the subjects with the GUI application and the localization task.

The questions section included two phases; questions with known direction and unknown distance and questions with unknown direction and unknown distance. This approach was used to evaluate how the knowledge of the direction of the received sound can modify the accuracy of the distance estimation. To avoid listener fatigue, a fixed subset of thirty out of sixty positions including every other position of each circle was used in both phases (marked in red in Fig.1). In each phase, each position was asked twice in a random order. Same positions were asked for all the listeners.

3.2.4 Questions phase 1

In the first phase, five speakers were shown on the screen with the same direction, but five levels of distance. The sound corresponding to one position was played and the listener was asked to click on the target speaker. The task considered for the subjects in this phase was only distance estimation as the direction was already given.

3.2.5 Questions phase 2

In the second phase, the listener was asked to choose the target speaker from all the sixty possible positions. Only one out of sixty positions was the correct answer where both the direction and the distance were distinguished correctly.

3.3 Listening test

Finally, the third step of this research was using the designed GUI to run the practical listening test for the subjects. Twenty test subjects (15 male and 5 female) who were all professional audio researchers participated in the test[6]. To play back the sound a pair of headphones[7] was used. The gathered data included the response of the individual test subject to each position for the both phases. The data analysis was performed by using MATLAB.

4 Results

To characterize the outcome of this study, dividing the main speaker pattern into four zones was beneficial as the results of three distinct directions in each zone tended to be similar (Fig. 4). Summarizing the final results with grouping the angles yielded better understanding of the effect of localization in four sides of the head on distance perception.

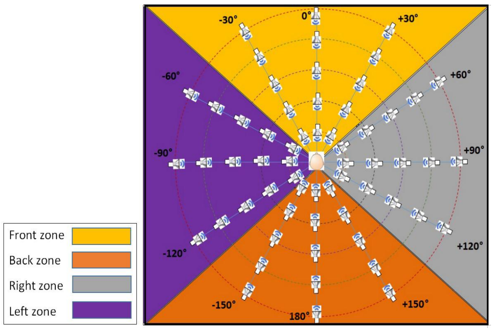

Figure 4. Recording set up divisions

4.1 Localization results

Fig. 5 and 6 show two sets of histograms to represent the listeners localization accuracy. This is to explain how many listeners distinguished the direction correctly from the second phase of questions. Fig. 5 shows the left and right zone results. From these histograms, it is evident that as the ITD became larger, the localization ability of the listeners

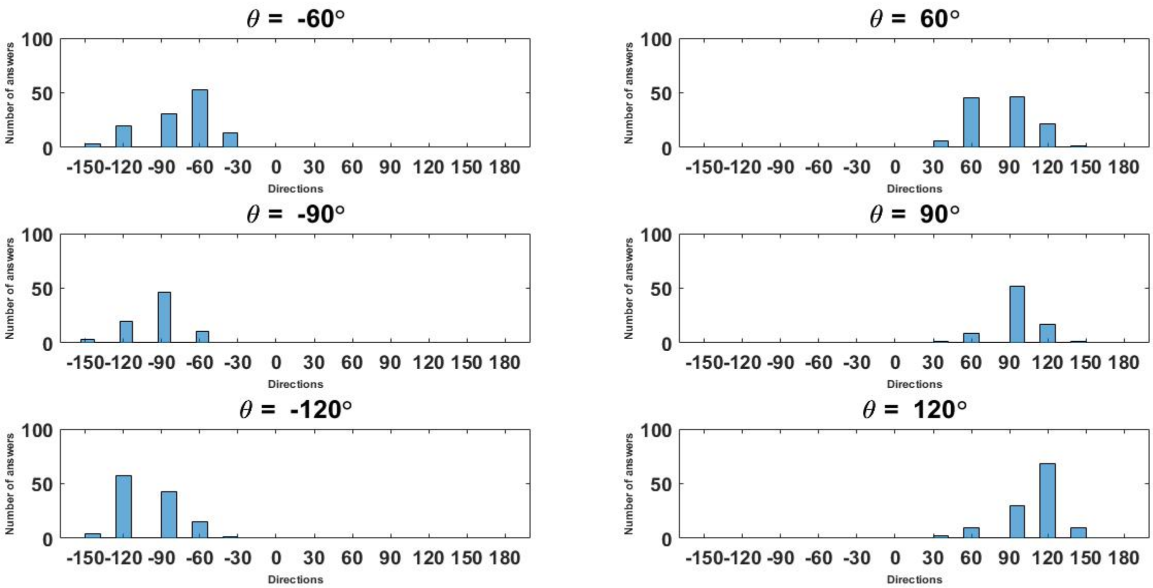

Figure 5. Pseudo-Normal distribution of data at left and right zone

[6] Location: University of California -San Diego - RML listening room - Atkinson Hall
[7] Sennheiser - HD 600

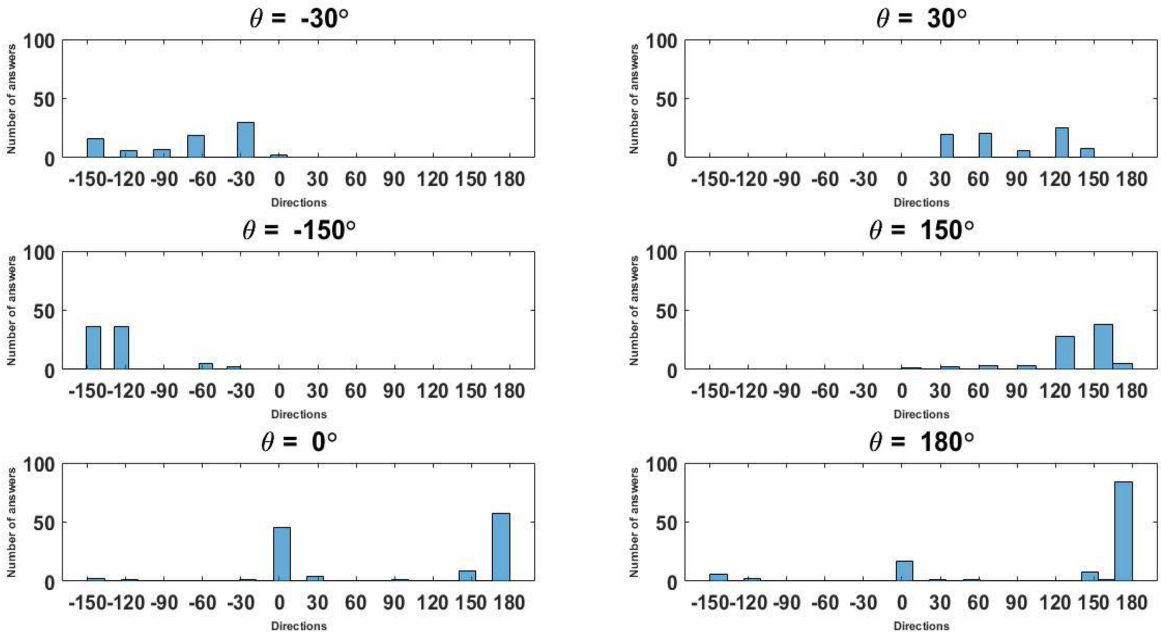

Figure 6. Non-Normal distribution of data at front and back zone

was better. These also illustrate a pseudo-normal distribution as opposed to the histograms shown in Fig. 6 which failed to show this. A wide range of answers were reported at -30 degrees and +30 degrees. Considering the results at 0 degrees and 180 degrees, the front-back confusion can be clearly seen and it seemed to be happening more frequently at 0 degrees. Using a dummy head for recordings might have intensified this effect.

4.2 Distance estimation results

To analyze the distance estimation results, two approaches were used; bar chart representation and chi-square test of independence [18] [19]. Bar chart was used to easily observe the relative proportion of correct distance estimation in each question phase. Chi-square test of independence was conducted to measure the significant difference between the outcome of the question phases. The significance level of the test (P-value) was set to 0.05. If the calculated P-value was equal or less than 0.05, the distance estimation accuracy was significantly improved having the direction of origin prior knowledge.

4.2.1 For the distance of 2 m

As shown in Fig. 7, the results of phase 1 at -60 degrees are 90% correct. Comparing the results at 60 degrees, the data obtained in phase 1 was 20% more precise (from 55% to 75%). The calculated P-value for this pair of angles was 0.0027 which represents a highly significant change in the estimation accuracy. This was the greatest difference between the outcome of phase 1 and 2 observed in this study. Bar charts corresponding to 120 degrees and -120 degrees showed 7% and 2% improvement respectively. The effect of the localization difficulty at 180 degrees on distance estimation can be seen in the bar charts. At this direction, no improvements were seen as the accuracy was reduced by 10% in phase 1. There was a 3% improvement in the results at 0 degrees which is not a significant change as the measured P-value was 0.7.

Generally speaking, the larger value of ITD and ILD at ±60 and ±120 located at left and right zones yielded better perception of direction and consequently distance in comparison with the directions in front and back zones.

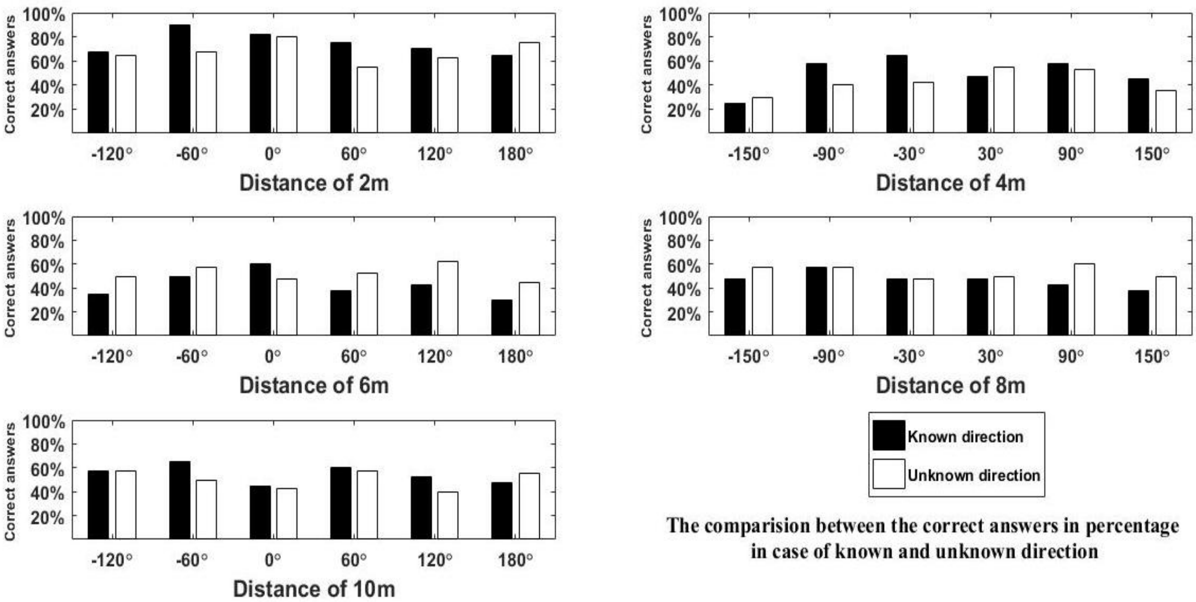

Figure 7. Number of correct answers in questions phase 1 and 2

In phase 1 at 2m, the correct answers varied from 65% (at 180 degrees) to 90% (at -60 degrees) which reported the most precise estimation in comparison with other levels of distance. This might be caused by the high sound intensity level at this distance. Another distance perception cue that was expected to play an important role at this distance was the highest Direct to Reverbarant energy Ratio. DRR was decreased with the distance increment and therefore less effective in larger distances.

4.2.2 For the distance of 4 m

At the distance of 4m, another set of angles was chosen. At -90 degrees, there was 17% growth in correct answers. The largest possible value of ITD and ILD at this angle could provided the listeners with more information about the direction of arrival and as a result, more precise distance perception. At 90 degrees, the results were improved by 4%. Moving on to the -150 degrees, it can be clearly seen that providing the direction did not raise the number of correct answers. According to the bar charts at 150 degrees, the correct answers were increased from 35% to 45%. To evaluate the significance of 10% improvements at this range, the P-value was calculated at this angle. The result was 0.36 which reported it was not a significant change. At ±30 degrees, the amount of the P-value was 0.34 which is higher than the significance level.

4.2.3 For the distance of 6 m

The third level of distance was shown on the next category of bar charts. The set of angles is similar to the set of angles shown at 2m. Bar charts corresponding to 180 degrees showed no improvements of data. This can be postulated as a consequence of correct answers having shrunk by 15% which was caused by the difficulty of localization in back zone. The graphs at 0 degrees highlighted a 12% increment in the correct answers (from 48% to 60%). The corresponding P-value was 0.26 which showed it was not a significant change. Other than the results of 0 degrees, bar charts at other angles presented here showed no improvements in the results of phase 1.

4.2.4 For the distance of 8 m

Moving on to the results of 8m, the graphs deal with the similar set of angles as 4m distance. The comparable results were shown here which indicated that the prior knowledge of the direction could not improve the estimation accuracy. To be specific, considering -150 degrees and +150 degrees, the number of correct answers were reduced by 9% and 12% respectively. At 90 degrees, the results of phase 1 are 17% less accurate than phase 2. At -30 degrees, +30 degrees and also -90 degrees, the correct answers did not grow although the subjects knew the direction of the origin.

4.2.5 For the distance of 10 m

Finally, the last set of bar charts illustrates the data of the 10 m distance which corresponds to the same set of angles as the distance of 2m and 6m. At 180 degrees, in phase 1, 48% of data is correct. In comparison with the outcome of phase 2 (55% correct answers) no improvements were seen. At 0 degrees and -120 degrees, according to the bar charts, the results of known and unknown directions were the same. At 120 degrees the results were raised by 13%, but not significantly (P-vale = 0.26). At -60 degrees, the outcome of the phase 1 is raised by 15%, but due to the measured P-value (0.1), it's not in the significant range. The results at 60 degrees were slightly changed by 3%.

5 Conclusion

This paper has evaluated the importance of considering the direction of incidence as a distance perception cue. The distance estimation results were described using bar chart representation and chi-square test of independence. The study showed that the direction of arrival effect varied from zone to zone. In the back zone, the prior knowledge of direction did not improve the distance estimation accuracy at any direction. In the front zone, at 0 degrees at 2 m and 10 m, the results of the known and unknown directions were so similar. Excluding the results for -30 degrees at 4 m, the results of 30 degrees and -30 degrees were not more precise in case of known direction in any distance level. In the right and left zones, the maximum improvement of the results of this research were reported at -60 degrees and +60 degrees at 2 m. The results at +90 and -90 degrees were improved at 4m. And finally, at +120 degrees and -120 degrees, the results were improved at 2 m.

6 Future work

The effect of distance perception on direction estimation accuracy should be also evaluated to create a general model of interaction between the localization in horizontal plane and distance perception. It should preferably include more angles and levels of distance.

Individual measured BRIR should be experimented for a reasonable number of subjects to do an assessment on the perception accuracy growth in the positions that are not estimated correctly in this study, such as front and back zones or higher levels of distance in left and right zone.

The large data set of this study could be subjected to more data modeling. For instance, the intra-subject reliability could be considered as a measure of subject's confidence in their answers which would suggest the sharpness of imaging.

So far, in this experiment, a stationary receiver - source scenario was investigated. However, to evaluate the effect of direction estimation on distance perception, a more dynamic scene can be investigated where either one or both the receiver and the source were in motion.

7 Acknowledgments

This research was funded by the Sonic Arts Research and Development group at Qualcomm Institute, the University of California San Diego Division of California Institute of Telecommunications and Information Technology (Calit2).

References

[1] Grantham, D. Wesley. "Spatial Hearing and Related Phenomena." *Hearing* (1995): 297, pp.323-328.

[2] Moore, Brian CJ. *An introduction to the psychology of hearing*. Brill, 2012, pp. 256-266.

[3] Kolarik, Andrew J., et al. "Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss." *Attention, Perception, & Psychophysics* 78.2 (2016), pp. 373-395.

[4] Shinn-Cunningham, B. G. (2000, December). Distance cues for virtual auditory space. In *Proceedings of the IEEE-PCM* (Vol. 2000, pp. 227-230).
[5] Zahorik, Pavel. "Auditory display of sound source distance." *Proc. Int. Conf. on Auditory Display*. 2002.
[6] C. W. Sheeline, An investigation of the effects of direct and reverberant signal interaction on auditory distance perception, Ph.D. Thesis, Stanford University, 1984.
[7] D. H. Mershon and E. King, "Intensity and reverberation as factors in the auditory perception of egocentric distance," Percept. Psychophys., vol. 18, pp. 409-415, 1975.
[8] D. H. Mershon and J. N. Bowers, "Absolute and relative cues for the auditory perception of egocentric distance," Perception, vol. 8, pp. 311-22, 1979.
[9] K. Fluitt, T. Mermagen and T. Letowski, "Auditory Perception in Open Field: Distance Estimation," Human Research and Engineering Directorate, ARL, MD, 2013.
[10] Lu, Yan-Chen, and Martin Cooke. "Binaural distance perception based on direct-to-reverberant energy ratio." *Proc. Int. Workshop Acoust. Echo Noise Control*. 2008.
[11] F. Everest and K. C. Pohlmann, Master Handbook of acoustics, fifth ed., McGraw hill, 2009, pp.56-60
[12] C. I. Cheng and G. H. Wakefield, "Introduction to Head-Related Transfer Functions (HRTF'S): Representation of HRTF'S in time, frequency and space," Audio Engineering Society Convention 107. Audio Engineering Society, Sep 1999.
[13] K. Bill , M. R. M. J. Michael and . M. Evangelos , "Auditory Perception and Spatial (3D) Auditory Systems," York University, Ontario, 2003.
[14] E. C. Sobel, "The locust's use of motion parallax to measure distance," Journal of Comparative Physiology A, pp. 579-588, 30 July 1990.
[15] Z. Chen and R.C. Maher, "Atmospheric Sound Propagation Considerations for the Birdstrike Project," progress report prepared for Advanced Acoustic Concepts, Inc., Hauppauge, NY, 10 pages, July 28, 2004.
[16] P. Zahorik, "Auditory Distance Perception, A Literature Review," University of Wisconsin - Madison, Madison, 1996.
[17] J. Blauert, "Spatial Hearing with one sound source", *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT press, 1997.
[18] Hogg, Robert V., and Allen T. Craig. *Introduction to mathematical statistics.(5"" edition)*. Upper Saddle River, New Jersey: Prentice Hall, 1995, pp. 269-278.
[19] Zibran, Minhaz Fahim. "Chi-squared test of independence." *Department of Computer Science, University of Calgary, Alberta, Canada* (2007).