Investigations on the Impact of Listener Movement on the Perception of Source Directivity in Virtual Acoustic Environments

Ulrike Sloma[1], Annika Neidhardt[1]
[1] Institute for Media Technology , 98693 Ilmenau, Germany, Email: ulrike.sloma@tu-ilmenau.de

Abstract

An authentic and plausible simulation of virtual acoustic environments requires, among many other aspects, a thorough understanding of how a listener perceives the directivity of a sound source. State-of-the-art VR applications allow the user to explore a virtual scenario while walking around freely. It is assumed that the translation of the listener in the scene contributes to the understanding of the scene and the sound source directivity. The aim of the research is to examine this impact of listener translations on the perception more closely. In this contribution a new test method to evaluate the perception using listener translation in virtual acoustic environments is presented. The experiment is conducted with binaural real-time auralization for tracked listener movements. A free-field scenario and a reverberant room are chosen as test conditions.

Introduction

This paper investigates an aspect of the research on the contribution and underlying cues of changes in listener position (translation) to the perception and understanding of virtual acoustic environments (VAEs). With the usage of dynamic binaural synthesis, complex acoustic surrounding scenarios can be auralized in a psychoacoustically correct way. Listeners are able to explore VAEs with rotation of their head and can freely walk inside the scenario. Tracking devices capture the position and orientation of the listener's head. The binaural rendering and playback over headphones is adapted corresponding to the tracked position. The possibility of continuous position change results in relative distance and angle changes between listener and sound source. This gives additional cues for sound source localisation, distance perception, perception of room acoustics, sound source characteristics and others. The overall goal is to determine the perceptual requirements for a stable, efficient and plausible simulation of dynamic VAEs. To achieve this overall goal a necessary step is to assess the impact of the listener translation on the perception in VAEs.

The study focusses on the assessment of the perception of sound source directivity in an interactive acoustic scenario. Several previous research activities addressed the influence of sound source directivities. San Martín et al. [1] showed that varying sound source directivities influence the measurements of room acoustic parameters. Further, Hoare et al. [2] stated that varying sound source directivities lead to significant perceptional effects in a virtual free field. Zotter et al. [3] claimed that it is hardly possible to differentiate between sound source directivities in a static listening position. However, if the listener are allowed to rotate the sound source they were able to align the directivity orientation towards their listening position. No studies investigating the movement of the listener in combination with varying sound source directivities have been found yet.

Based on the previous research the assumption is that moving along or around a virtual sound source enables a listener to perceive and distinguish different sound source directivities in an analogous way the rotation of a sound source does.

Methodology

An evaluation approach is developed to get an insight if there is an impact of listener movement on the perception of sound source directivities. Three research questions are formulated for the investigations.

1. Can listeners distinguish directivities of sound sources in a virtual acoustic environment...
    a) with head rotation only?
    b) with head rotation and change in listener position?
2. Do the room acoustics influence this perception?
3. Can a listener learn directivity patterns while exploring a room with movement and apply this patterns in a static listening situation?

To answer these questions a test methodology for comparison of static listening and listening with positional change is designed. For sound sources with differing directivities the ability to distinguish between those is investigated.

Simulations, test-stimuli and reproduction

The used Binaural Room Impulse Responses (BRIRs) for the representation of the VAE are calculated with the Matlab based simulation toolbox MCRoomSim [4]. Two shoe-box rooms with the same dimensions ($8.40 \times 7.60 \times 2.80\,\mathrm{m}^3$), but different reflection and absorption coefficients are defined. One is a strongly absorbent room, with nearly anechoic characteristics ($\mathrm{T}_{60} = 0.05\,\mathrm{s}$). This room is called free-field in the following. The second room has stronger reflections ($\mathrm{T}_{60} = 0.69\,\mathrm{s}$) and is called reverberant room. In the rooms a sound source with varying directivity is placed at $\mathrm{x} = 3.00\,\mathrm{m}$, $\mathrm{y} = 3.15\,\mathrm{m}$ and $\mathrm{z} = 1.55\,\mathrm{m}$. Two directivity characteristics are used, which are already available in MCRoomSim. One is an omni-

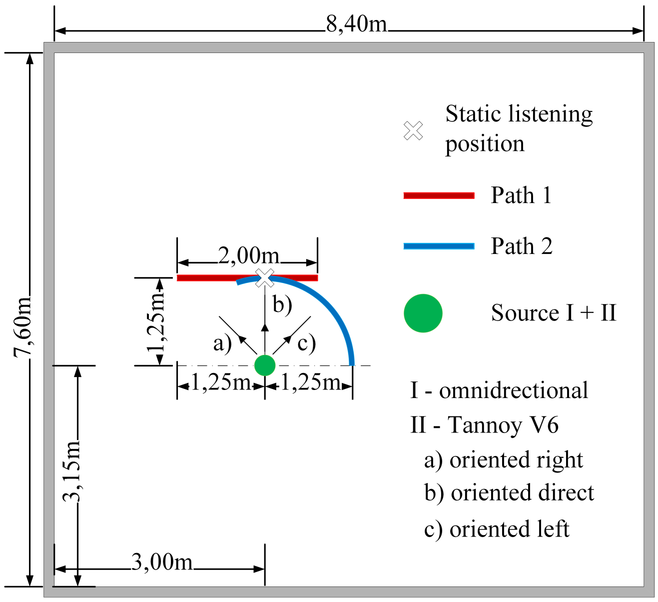

Figure 1: Sketch of the simulated room dimensions including the static listening position, the walking paths, sound source position and orientations.

directional sound source and the other a sound source with directional characteristic according to a TannoyV6 loudspeaker. The loudspeaker can have three orientations, 45 ° right, 45 ° left and direct as shown in figure 1. The BRIRs are calculated with a positional resolution of 25 cm for two walking paths. First a straight line, path 1, with varying listening distances to the sound source and second a quarter circle, path 2, with equal listening distances to the sound source.

For the binaural representation the Neumann KU100 Dataset of Head Related Transfer Functions (HRTFs) is used [5]. A full 360 ° rotation with an angular resolution of 4 ° is calculated for each simulation position.

The calculated BRIRs are finally auralized with a pop song, according to the position and listening orientation of the listener. The tracking is done with an HTC Vive and the respective BRIRs are convolved in real-time with the audio using the PyBinSim python package [6]. STAX headphones are used for the presentation of the corresponding auralized signal. Visual feedback of the listener position is given through the integrated camera, which displays a blurred image of the surrounding real room.

Listening test design

The test is conducted in a listening lab according to ITU R BS.1116 [7]. The listeners are able to move freely along the pre-defined paths and explore the presented acoustic scenarios.

The test procedure is structured in four phases. No training session is included, since the initial ability of the listener to distinguish sound source directivities should be assessed.

Phase 1: Static 1 - listening in one static position with head rotation only

Phase 2: Path 1 - exploring the acoustic scene with translation and head rotation on path 1

Phase 3: Path 2 - exploring the acoustic scene with translation and head rotation on path 2

Phase 4: Static 2 - repetition of phase 1

In phase 1 and 4 all described variations, the combination of all 4 directivities and both rooms, are presented to the listener. The listening position is marked by an × in figure 1. Phase 2 and 3 only include the omnidirectional directivity and the direct sources oriented towards the path. All stimuli are evaluated twice and in randomized order. The two paths are evaluated with changing order by the participants.

The listener is asked to state for each stimuli which sound source directivity and if applicable with which orientation is perceived relative to the direct listening position. The overall possible options are explained in an introduction before the test. The answers are written down by the test coordinator who is inside the listening lab during the whole test procedure.

Results

In the study 20 listeners took part, three female and 17 male, aged between 23 and 40 with an average at 30. All participants are dealing with audio and acoustics in their profession or leisure time. So the most are not expert listeners, but they have a certain experience with similar tasks.

For test phase 1, static listening 1, the longest evaluation time can be observed. With an average of 12.1 minutes this time is significantly longer than the duration of phase 4, the repetition of phase 1, with an average of 9 minutes. The averaged durations for phase 2, 9.05 minutes , and phase 3, 9.45 minutes, do not differ. Figure 2 illustrates the analysis of the listening times for all participants and phases.

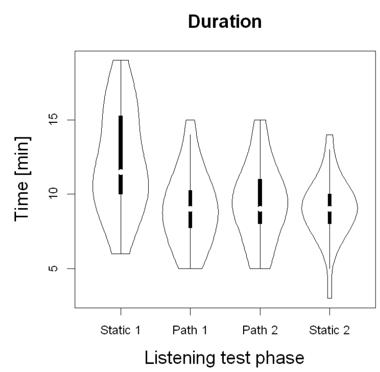

Figure 2: Duration of the four phases.

All participants answered correctly for minimum 22 and maximum 41 out of 56 total stimuli. The average is 29.35 correct answers. The probability that a listener is only guessing is below 5 % for at least 19 correct answers.

Static listening (phase 1 and 4)

First the listening in a static position with allowed head rotation is analysed. A comparison of the evaluations before and after listening with positional changes is conducted. Figure 3 displays selected diagrams of the results. A description of the diagrams is added in the caption.

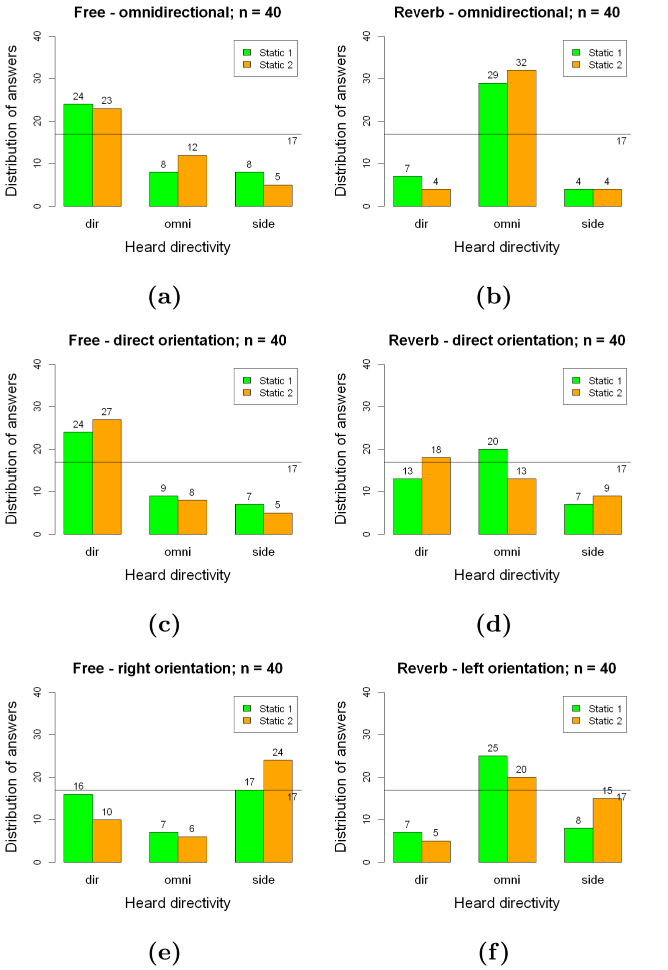

Figure 3: Selected results for static 1 and static 2. Shown are the distributions of the answers from all participants for each, twice presented, stimuli (40 answers total). For easier interpretation the answers for left and right are summarized as "side". The green bar marks phase 1, static listening 1, and the orange bar marks phase 4, static listening 2. For values above the horizontal line the guessing probability is below 5 %.

The results for the simulated omnidirectional source directivity in the free-field condition, see diagram 3a, and the reverberant room, see diagram 3b, show that it is initially not possible to perceive and assign the intended source directivity. In free-field the stimuli are stated as directed towards the listening position and in the reverberant room as omnidirectional. It is assumed that the participants decided according to the respective room acoustic characteristics. In diagram 3d the directed sound source with orientation towards the listening position in the reverberant room is shown. In the first static listening phase the previously mentioned assignment to the omnidirectional characteristic can be seen. Some participants are able to state that the sound is oriented with direct characteristic towards their listening position after the phases with listener movement. The effect cannot be seen in the free-field condition 3c and the omnidirectional sound source in both rooms.

For the directed sound sources with orientation towards the both sides (left and right) several correct answers are recognized, especially in the free-field. In diagram 3e the distribution for the right oriented source in free-field and in diagram 3f for the left oriented source in reverberant room is displayed. Again, the effect of the assignment of sources in the reverberant room to be omnidirectional can be seen. The number of answers towards the side tend to increase after the phases with listener translation and the possibility to explore the scene, which might indicate a slight learning process.

Listening with positional changes (phase 2 and 3)

To analyse the results of phases 2 and 3 the answers of the participants are redefined to be able to compare both paths. Afterwards, the four possible answers, shown in figure 4, are omnidirectional, towards the middle of the path, towards the right side of the path or other, if something else is stated by the listener. Figure 5 displays selected diagrams of the results. A description of the diagrams is added in the caption.

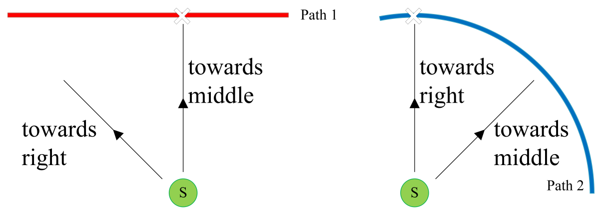

Figure 4: Redefinition of listener answers for comparative analysis of path 1 and path 2.

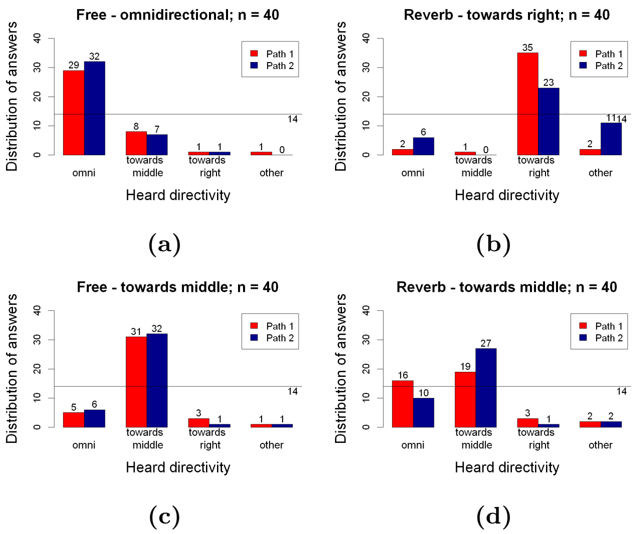

Figure 5: Selected results for path 1 and path 2. Shown are the distributions of the answers from all participants for each, twice presented, stimuli (40 answers total). The red bar marks phase 2, path 1, and the blue bar marks phase 3, path 2. For values above the horizontal line the guessing probability is below 5 %.

Diagrams 5a and 5b display the distribution for the simulated omnidirectional source in the free-field condition and the source oriented towards the right of the paths in the reverberant room. For all omnidirectional and towards right oriented stimuli the participants are able to clearly distinguish between the intended directivities in both simulated room conditions and for both paths. The directed source with orientation towards the middle of the path in free-field can also clearly be stated, as shown in diagram 5c. Diagram 5d shows the corresponding results for the reverberant room. Here confusions with the omnidirectional sound source directivity occurred especially for path 1 with varying distances to the sound source.

Decision criteria

Upon completion of the test process the participants are asked about their criteria to differentiate between the sound source directivities. When listening at fixed position and allowed head rotation, the most participants are undecided and try to listen to the reverberations and reflections of the presented room. Some mention a procedure of exclusion but without any further description. The experiences of the phases with positional changes are used, or tried to be used, for evaluations in phase 4. In the test phases with translation on the pre-defined paths the main criteria are changes in timbre, dull versus bright sound experience, or changes in loudness. A few people try to listen explicitly to the instruments, the understandability of the singer or the direct to reverberant ratio.

Summary and conclusions

This contribution presents a study on the ability to understand and distinguish sound source directivities with and without positional change of a listener in VAEs. The goal is to prove that head rotation only is not sufficient to perceive the directivity of a simulated sound source and that additional translation and exploration of a scenario, for instance with pre-defined walking paths, is helpful for a comprehensive understanding of the VAE.

With the analysed results it can be concluded, that static listening with allowed head rotation provides not enough cues and information of the acoustic scene to decide whether a perceived sound source is omnidirectional or directed with a specific orientation. Initial the listeners match reverberant room acoustics with omnidirectional sources and near anechoic room acoustics with directional source directivities. Perceptive changes of sound sources oriented towards the side can be observed. In contrast, if listeners are enabled to walk inside the room, enough cues, like changes in timbre or loudness, are provided to distinguish between varying sound source directivities. A learning effect cannot be seen with statistical certainty in the data of phase 4. More relevant effects are expected with more training and self-exploration of the acoustic scenario. Distance changes to the sound source when walking, path 1, influence the assessment of the sound sources towards the middle of the path in the reverberant environment. It is expected that stronger reverberations enhance this effect.

The research questions from the introduction are answered as follows: Listeners are not able to distinguish directivities of sound sources when only head rotation is allowed. For listening with positional changes a distinction is possible. The room acoustics have a high influence on the perception in static listening, but not in translational listening. A learning process cannot be shown with certainty.

Overall, the results confirm the assumption that there is a considerable contribution of listener translation to the understanding of sound source directivities and therefore the respective VAE. In further research these results will be validated and it will be investigated which parameters and cues are used to distinguish between sound source directivities.

Acknowledgement

The authors like to thank all the participants who were taking part in the listening test. This work is supported by a grant of the Deutsche Forschungsgemeinschaft (Grant BR 1333/18-1).

References

[1] R. San Martín, I. Witew, M. Arana, and M. Vorlaender. “Influence of the Source Orientation on the Measurement of Acoustic Parameters”. In: *Acta Acustica united with Acustica*. Vol. 93. May 2007, pp. 387–397.

[2] S. Hoare, A. Southern, and D. Murphy. “Study of the Effect of Source Directivity on the Perception of Sound in a Virtual Free-Field”. In: *Audio Eng. Soc. Conv. 128*. 2010.

[3] F. Zotter, M. Frank, A. Fuchs, and D. Rudrich. “Preliminary study on the perception of orientation-changing directional sound sources in rooms”. In: *Proc. of Forum Acusticum, Kraków*. 2014.

[4] A. Wabnitz, N. Epain, C. Jin, and A. Van Schaik. “Room acoustics simulation for multichannel microphone arrays”. In: *Proc. of the Int. Symp. on Room Acoust.* Citeseer. 2010, pp. 1–6. URL: `https://www.ee.usyd.edu.au/carlab/mcroomsim.htm`.

[5] B. Bernschütz. “A spherical far field HRIR/HRTF compilation of the Neumann KU 100”. In: *Proc. of the 40th Italian (AIA) Annu. Conf. on Acoust. and the 39th German Annu. Conf. on Acoust. (DAGA)*. 2013, p. 29. URL: `http://audiogroup.web.th-koeln.de/ku100hrir.html`.

[6] A. Neidhardt, F. Klein, N. Knoop, and T. Köllmer. “Flexible Python Tool for Dynamic Binaural Synthesis Applications”. In: *Audio Eng. Soc. Conv. 142*. 2017. URL: `https://github.com/pyBinSim/pyBinSim`.

[7] R. ITU-R. “BS. 1116-1: Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems”. In: *Int. Telecomm. Union, Geneva Std* (1997).