DOI: 10.22032/dbt.39967
DOI (proceedings): 10.22032/dbt.39936

Abstract Reviewed Paper at ICSA 2019

Presented * by VDT.

Investigation on spatial auditory perception using non-uniform spatial distribution of binaural room impulse responses

Stephan Werner[1], Florian Klein[1], and Georg Götz[2]
[1] *Technische Universität Ilmenau, Germany, Email: {stephan.werner; florian.klein}@tu-ilmenau.de*
[2] *Aalto University, Finnland, Email: georg.gotz@aalto.fi*

Abstract

For spatial audio reproduction in the context of virtual and augmented reality, a position-dynamic binaural synthesis can be used to reproduce the ear signals for a moving listener. A set of binaural room impulse responses (BRIRs) is required for each possible position of the listener in the room. The required spatial resolution of the BRIR positions can be estimated by spatial auditory perception thresholds. If the resolution is too low, jumps in perception of direction and distance and coloration effects occur. This contribution presents an evaluation of spatial audio quality using different spatial resolutions of the position of the used BRIRs. The evaluation is performed with a moving listener. The test persons evaluate any abnormalities in the spatial audio quality. The result is a comparison of the quality and the spatial resolution of the various conditions used.

1. Introduction

Existing audio systems can reproduce spatial audio in a way that artificial and real audio objects are perceived as plausible audible events in a virtual and/or augmented environment [3]. An auditory illusion of a spatial acoustic environment can be created with the help of existing position-dynamic binaural synthesis systems [2]. The occurrence of such a plausible auditory illusion depends on an adequate technical realization and on several context dependent quality parameters like congruence between synthesized scene and the listening environment or individualization of the technical system for example.

This paper examines spatial auditory perception thresholds using a position-dynamic binaural synthesis. The binaural transfer functions are provided for discrete positions in the room. The local area in which one set of BRIRs is used is referred to as cell. The size of the cell directly influences the direction and distance errors of the reproduction caused by it. The discretization of the room is therefore determined by the perceived minimum direction and distance change between the position of the listener and the source. This allows the creation of a perceptually motivated BRIR grid which needs less BRIRs than a uniform shaped grid.

2. Binaural Synthesis System

The reproduction of an audio object in a reverberant room can be realized by using BRIRs. The audio signal of the source is convolved with the BRIRs for the left and the right ear and for the current source-receiver position and head orientation of the listener. A change of position and/or head pose requires a new pair of BRIRs. The position and pose changes are continuously measured by a tracking system and made available to the BRIR selection. In this contribution, a QualiSys motion capturing system is used to track the horizontal orientation and the x-y coordinates of the listeners‘ head. Headphones are mostly used as playback devices. An open or extra-aural headphone additionally enables an acoustic recognition of the real environment. The headphones must be equalized for correct reproduction of the binaural ear signals. For this contribution a KEMAR head and torso simulator is used for the BRIR recordings and an equalized Beyerdynamic DT-1990pro headphone is used for playback. The inverse of the headphone transfer function is calculated by a least-square method with minimum phase inversion [13].

The needed BRIRs can be calculated by room-acoustic simulations or by measurements of real sound sources in a real room. A comprehensive synthesis of an auditory scene with a variety of sound sources, room acoustics, and movements of sources and receiver requires a high number of BRIRs. A minimization of this number while simultaneously maintaining a high perceived quality is desirable here.

Several approaches are available to reduce the amount of measurements for binaural synthesis. These are for example room acoustic simulations [12, 20] or methods which use head-related transfer functions and directional dependent or independent room impulse responses [6, 11]. In addition to these approaches the required BRIRs can be synthesized from the measurement of just one BRIR data set from only one position in the listening room by changing acoustic parameters like initial time delay gap, energy decay, and direct to reverberant energy ratio. These methods are applied for this contribution. Details can be found in the references [16, 18].

A basic feature of the used approach is the use of synthesized BRIR data sets for discrete positions in the room. These discrete areas can either have a uniform distribution (grids with rectangular grid areas/cells) or a non-uniform distribution of the single grid cells. Figure 1 illustrates an uniform grid. The listener can move within an area of max. 4 m x 4 m. The resolution of the binaural synthesis for translation is 0.25 m using an uniform grid of squares.

2.1. Basic Arrangement

Figure 1 illustrates the walkable area and the location of the possible audio object positions. The center speaker (C) is positioned at 0° and 3.5 m from the midpoint of an assumed circle. A left and a right speaker (L, R) is positioned at +/- 30°, another left and right speaker (SL, SR) is placed at +/- 120°. For the recording of the BRIRs, loudspeakers of the type Geithain MO2 were used. They were arranged at ear level of the used KEMAR artificial head. Their orientation was towards the midpoint of the area. Only the middle position of the area was measured. The other grid positions are synthesized as described above.

The concentric black line in figure 1 shows the path to be followed in the listening test by the listeners. More details about this can be found in section 3.2 Test Procedure.

2.2. Non-Uniform Grids

The use of a uniform grid as shown in figure 1 does not take perceptive inaccuracies in localization [7] and distance perception [1, 14] into account. Figure 2 shows the principle of localization blur and distance blur in the horizontal plane. If a fixed localization blur or minimum audible angle (e.g. 5°) is assumed, the density of the cells (in the sense of the

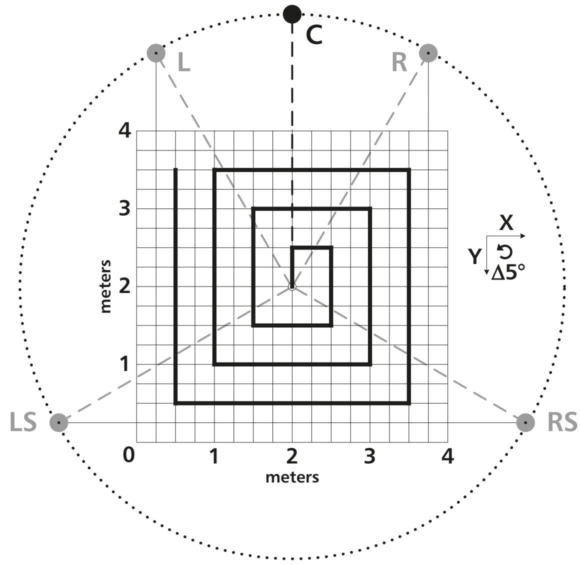

Fig. 1: Schematic view of the walkable area with the possible audio object positions. Concentric black line indicates the walking path.

width of the cells) should have to increase at small distances to the source and decrease at larger distances. The situation is similar with regard to distance blur. For small distances the density (in the sense of the length of the cells) should be higher than for large distances. This approach leads to the non-uniform cells which can effectively save on the number of BRIRs required without causing increased errors in direction and distance perception.

In each individual grid cell, the BRIR set inside only represents the distance from the sound source to the center of the cell. If the listener is not in the center of the cell, the distance and direction cues no longer corresponds to the intended object position. The transition between the cells leads to a maximum deviation.

Figure 3 shows the four non-uniform grids used on the basis of a Voronoi network. Grid 1 (Figure 3 top left) is the finest grid and Grid 4 (Figure 3 bottom right) has the lowest resolution. The main parameters for generating the grids are the maximum allowed angle error α and the maximum allowed relative distance change d. For the grids were used:

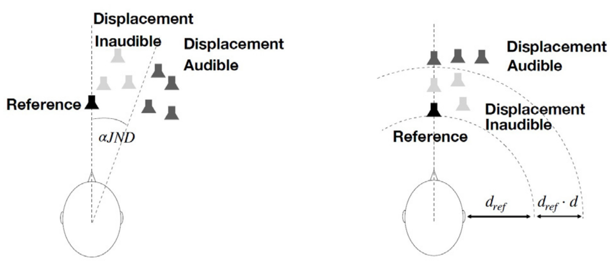

Fig. 2: Localization blur (left) and distance blur (right) in the horizontal plane. A displacement of the reference (black) by an angle or distance below the threshold is inaudible (light grey), whereas a bigger displacement is audible (dark grey). (from [4])

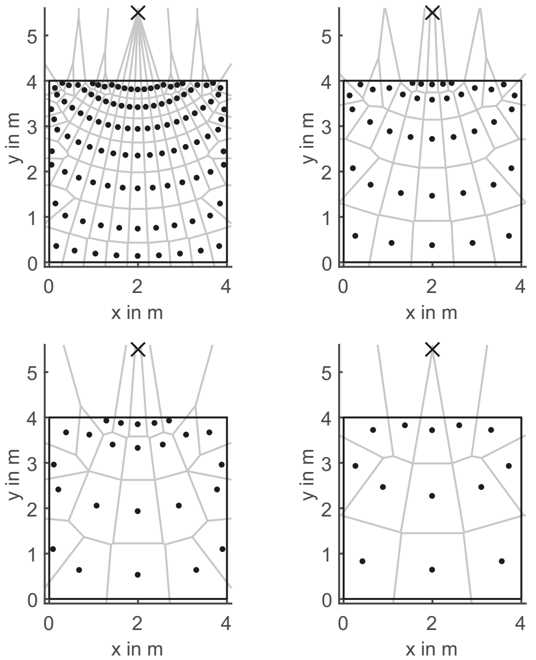

Fig. 3: Non-uniform grids used in the listening test; top left... grid 1, top right... grid 2, bottom left... grid 3, bottom right... grid 4.

Grid 1 $\alpha = 5°$, $d = 0.25$; Grid 2 $\alpha = 10°$, $d = 0.5$; Grid 3 $\alpha = 15°$, $d = 0.75$; Grid 4 $\alpha = 20°$, $d = 1.0$. The distance parameter for Grid 1 is based on results from Spagnol et al. [15] where distance blurs of $d = 0.25$ were found. The angle parameters are estimates which have been collected from the literature. It is assumed that the localization blur range is from $1°$ to $10°$. The $\alpha = 5°$ for Grid 1 is selected on the one hand because it is in the middle of the indicated range and on the other hand it corresponds to the angular resolution of the BRIRs for head rotation. The perceptive effect of a small angle for localization blur would be masked in an evaluation by the larger angle for the head rotation.

For a detailed description of the procedure for creating non-uniform grids, we refer to the work of Georg Götz and Samaneh Kamandi [4]. This work can be requested from the authors of this contribution.

For each cell, a BRIR data set is stored for the left and right ear, which has a horizontal resolution of 5° for the head rotation. Therefore 2x72 BRIRs are stored per cell and for each sound source. Table 1 shows the number of cells and number of BRIRs for the different grids. Compared to an assumed uniform grid with the same angle resolution of 5°, a grid resolution of 0.25 m, and for one source, as shown in Figure 1, there are high reductions in the number of BRIRs required. The reduction as a ratio of the number of BRIRs required for the uniform grid to the number of the sparsest non-uniform grid is 19.7 (see Table 1).

The centers of the individual cells are marked by dots in figure 3. Only at these points the BRIRs reproduce the correct direction and distance information. Outside the center, errors occur due to a shift in distance and direction. This effect is evaluated in the listening tests.

Tab. 1: Overview of the number of cells and BRIRs for the different grids. The ratio indicates the saving of BRIRs to the uniform grid.

Grid	number of cells	number of BRIRs	ratio
uniform	256	36,864	-
1	111	15,984	2.3
2	39	5,616	6.6
3	24	3,456	11.0
4	13	1,872	19.7

3. Listening Test

The listening test is intended to investigate various non-uniform grids with regard to spatial auditory perception. The aim is to find out what effect a reduction in the number of BRIRs has. Two evaluation approaches are used for this purpose.

3.1. Test Design

The listening test is divided into two parts. In the first part, unanticipated events should be detected while the listener is walking around. These events should refer to artifacts in the spatial listening which may lead to an implausible scene perception. This can be for example: object extension, distance, localization, envelopment, externalization, spaciousness, and timbre.

The second part of the test takes place each time a test condition, and therefore a path, has been completely passed through. The test person evaluates various quality features in a questionnaire (see subsection Quality Evaluation 3.1.2). Of course, this does not make the single unanticipated event assessable, but at least the quality impression of the single test conditions.

3.1.1. Detection of Unanticipated Events

The test persons should confirm a detected unanticipated event while walking around with the help of a radio button. The time and especially the x-y coordinate in the grid is recorded as information using the tracking system for the binaural synthesis. The button itself is realized by a presenter for slide presentations via a radio link and queried with the help of a Python script.

The test person is allowed to explore the local area more precisely by small forward and backward movements and head rotations when an unanticipated event is detected. This should allow a higher reliability in the determination of the position.

3.1.2. Quality Evaluation

The evaluation of perception is performed for different quality features. In addition to single features, this also includes the rating of the overall impression as a kind of overall quality. The individual quality features include localization, externalization, and timbre perception. The query usually takes place on quasi-continuous scales with the negative value at ”0” and the positive value at ”100”. Externalization is rated on a category scale.

Overall Impression OI - The overall impression should capture the individual impression and the perception of the overall quality of what is heard by the listeners. No specifications were made with regard to possible underlying single quality features. The survey of the overall impression was always carried out at the beginning of the survey for each walking path in order to minimize the influence of the evaluation of single quality features. The following question had to be answered: ”How would you rate the quality of experience of this system?” A quasi-continuous rating scale from ”0-poor” to ”100-very good” was used.

Localization Ability LA - The localization ability is intended to test the ability of the listener to localize the auditory event. A high localization ability exists when the auditory event can be clearly assigned to a direction. If the directional information of the auditory event is diffuse or not localizable, a low rating should be given. The following question had to be answered: ”How would you rate the ability to localize the audio object?” The quasi-continuous scale used ranges from ”0-poor” to ”100-very good”.

Localization Stability LS - The use of different grids with different spatial resolution can lead to jumps in the perception of the object position. The movement of the listener through the individual grid cells causes more or less pronounced abrupt changes in the reproduction of directional and distance cues. The following question had to be answered: ”How would you rate that the audio object stays at a fixed position?” The quasi-continuous scale used ranges from ”0-poor” to ”100-very good”.

Coloration CO - The feature coloration aims at the evaluation of the hearing perception, which cannot be described by directional hearing, perception of externalization or spatial stability of the auditory event. The underlying perceptual feature of coloration is timbre. Timbre can be defines as ”that attribute of sensation in terms of which a listener can judge that two steady complex tones having the same loudness, pitch and duration are dissimilar” [9]. In the present case, timbre is defined as the difference in perception of the audio signal at a detected abnormality and the otherwise perceived audio. This is referred to as coloration. The audio signal of the abnormality is evaluated as colored compared to the remaining audio shortly before it. The following question had to be answered: ”How would you rate the coloration of the audio during walking?” The quasi-continuous scale is in a range from ”0-strong coloration” to ”100-no coloration”.

Externalization - The externalization of auditory events describes the perception of the location of the event in the head or outside the head of the listener [5, 10]. Externalization is a crucial feature to reach a plausible spatial auditory illusion with binaural headphone systems [10, 17]. The dichotomous quality feature is counted as the index of the ratings on a three-point scale. In additions to the characteristics “in-head” and “outside the head” a transition point “outside but close to the head” is used. This scale is motivated by the individual mapping to a scale of the percept of externalization for every test person. Only the scale point “outside the head” is counted as an externalized auditory event in further analysis. We define the perception of an event very close to the head or ears as in-head-localized or non-externalized. We suppose that this conservative approach maps the ratings in a reliable way referred to the resynthesis of the real loudspeakers with their positions in the room. The goal is to minimize the confusion between distance perception and externalization for closer distances. The following questions had to be answered: ”How would you rate the externalization of the audio object?”. The following scale is used ”1=in-head”, ”2=close to the head”, ”3=in the room”, ”4=at the intended distance”.

3.2. Test Procedure

The test was divided into several phases. The first phase comprises a written and oral introduction to the test environment, the assessment methodology and the quality characteristics to be assessed. The second phase is the familiarization phase with the position-dynamic binaural synthesis system. The headphones were placed on the listeners and were not removed until the end of the entire test. This should make it possible to get used to the headphones. The test persons should continue to move freely within the accessible area. Two audio scenes were synthesized binaurally.

The **first phase** consists of different male and female speakers who are placed at the five loudspeaker positions. The scene starts with one speaker until all five speakers are active at the same time. The intention of this scene is to capture individual sound object positions, to get used to a complex scene and to promote the active movement of the listener by specifically listening to individual sound sources, also by moving towards the sources.

The **second phase** is an excerpt from a radio play in which there are discrete audio objects and an enveloping and ambient sound-scape. In contrast to the first scene, this scenery is realized by multi-channel stereo panning onto the five loudspeakers. The intention of this scene is to get used to a complex environment and to listen to the content without consciously considering the technical realization. A uniform grid with a spatial resolution (spacing of the grid cells) of 0.25 m and an angular resolution of the horizontal head rotation of 5° was used for playback. This resolution leads to slight perceptible artifacts in critical hearing in the form of localization jumps of the audio object position in translational and rotational movements. A technical and perceptive proof of function of such a system is to be taken from the references [8, 18, 19].

The **third and final phase** includes the evaluation of the individual grid conditions. At the beginning, a section of an audio book is presented to the test persons. The intended source position is the center position of the arrangement used. The test persons should follow the path shown in figure 1 in order to cover an area as large as possible uniformly. The concentric path starts in the middle of the walkable area and leads to the outside. Its length is $24m$ for one walking direction. The path had to be walked forward and backward. According to the instructions the test persons had to mark unanticipated events by pressing a radio button. At first the worst resolution grid was presented. This should ensure that abnormalities are also found. The different test conditions (different grids and audio signals) are then presented in random order for each test person. After the path for each condition had expired, the individual quality characteristics were evaluated in a questionnaire. After this the next test condition is presented. The test persons were instructed to position their heads in the walking direction. However, head movement was explicitly allowed. The test persons were also instructed to move in a normal to slow walking speed. After the evaluation of the last test condition, the test persons had the opportunity to make further comments and remarks on the test. The whole procedure took about one hour.

4. Ratings

The evaluations of 21 out of 22 test persons with a mean age of 29.1 years (standard deviation 8.2 years) were used for the evaluation. Fifteen of the five women and 17 men were experienced in listening tests and ten persons had experiences with binaural synthesis systems. Experience is defined as participation in at least two listening tests or the listening to a binaural synthesis at least twice. The ratings of one test person were omitted because even for the playback with the worst grid, ratings were always given on the scales of more than 85 scale points and thus clearly outside the average of the ratings of the other test persons. No further post-screening of the assessments was carried out. The authors are not aware of any method that allows a reliable and comprehensible evaluation of spatial hearing quality assessments using a virtual hearing instrument. The test itself did not contained any evaluations of real sound sources that might have allowed this.

Table 2 shows the time required for the path to walk forward and backward. The average speed corresponds to a slow walking speed. The standard deviation as well as the max and min values indicate a certain variance in walking speed between the test persons. The mean walking speed was at $0.4m/s$, with a length of the whole path of $48m$.

Tab. 2: Duration of walking of the test persons to complete the path (forward and backward); times are in minutes; N indicates the number of the single walks.

N	mean	standard deviation	max	min
160	02:01	00:22	03:02	01:17

4.1. Detection of Unanticipated Events

Figure 4 shows the total number of ratings for the detection of unanticipated events during listening. The number is proportional to the grid resolution. There is an almost linear increase with relation to the grids, which indicates an uniform selection of the grids with relation to its parameters. There is a higher number of ratings for the music signal than for the speech signal. This may indicate that the music signal was rated as a more critical signal. The test subjects' statements at the end of the test also indicate this.

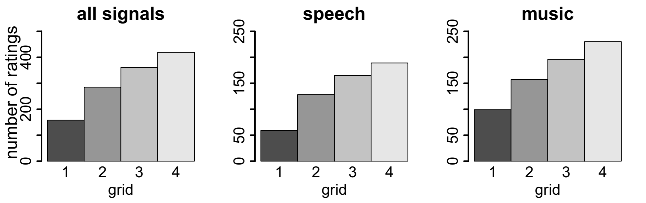

Fig. 4: Absolute number of ratings of unanticipated events for all grids and audio signals.

The figures 5 and 6 show the frequency of perceived unanticipated events as a heat map while walking around. The side plots show the sum of the frequencies for the x and y axis as histograms. The gray lines shown in the heat map illustrate the position of the respective grid cells. Only the ratings for the most highly resolved grid (grid 1, figure 5) and the least resolved grid (grid 4, figure 6) are shown. The ratings for the other two grids lie between these two.

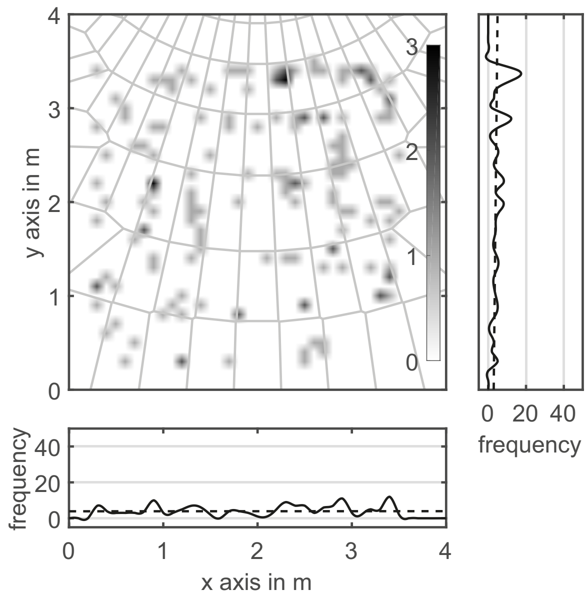

Fig. 5: Ratings of unanticipated events as heat-map for Grid1 ($\alpha = 5°$;$d = 0.25$). Side plots show the sum of the ratings over the x or y axis as a histogram. Dashed lines indicate regression line. Grey lines illustrate the position of the respective grid cells.

When the audio signals are played back using Grid 1 (Figure 5), the number of detected abnormalities is significantly lower compared to grid 4 (Figure 6). For all grids there is tendentially a symmetrical distribution of the abnormalities along the x-axis. For the y-axis there is a decrease in the number of abnormalities with increasing distance to the source. This effect increases with the use of more coarse grids.

4.2. Quality Evaluation

Figure 7 shows the ratings of perceived externalization of auditory events for the various grids and audio signals. When using the grid with the highest resolution (Grid 1), high externalization indices close to 1 are visible. The indices decrease continuously but slightly to a value of around 0.75

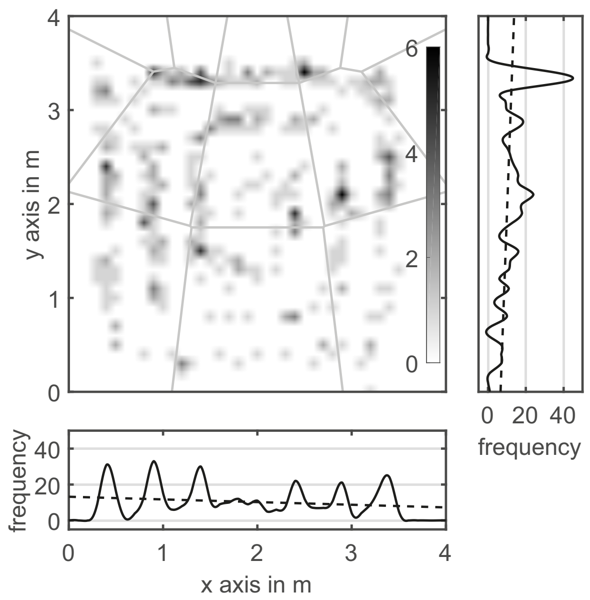

Fig. 6: Ratings of unanticipated events as heat-map for grid 4 ($\alpha = 20°$;$d = 1.0$). Side plots show the sum of the ratings over the x or y axis as a histogram. Dashed lines indicate regression line. Grey lines illustrate the position of the respective grid cells.

when less high-resolution grids are used. It is assumed that the localization and distance errors lead to a reduction of the externalization due to a reduced plausibility of the perception. With regard to the audio signals used, there are no influences on the externalization observed.

Figures 8 and 9 show the ratings for the individual quality features and for the overall impression for the music and the speech signal. It can be said that Grid 1 achieved the highest ratings. There is a tendency for lower valuations to be given for the less resolved Grids. The interquartile distances tend to rise slightly for Grid 3 and 4. This indicates that the test persons are less in agreement if grids with lower resolutions are used.

Overall it must be noted that relatively high quality ratings are given both for the overall impression and for the individual quality features in view of the savings achieved in the number of required BRIRs. However, it must also be made clear that the perceived quality also depends on the type of application. For example, a test scenario with a direct comparison of real and virtual sources with the intention of testing for

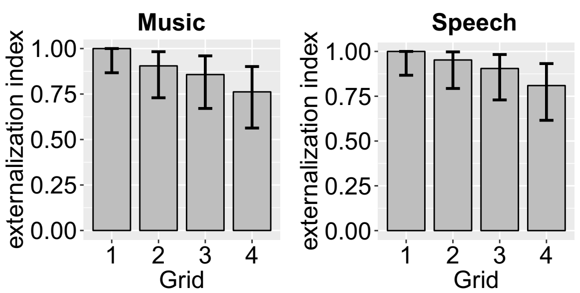

Fig. 7: Externalization as index with 95% conf. interval; left: music, right:speech.

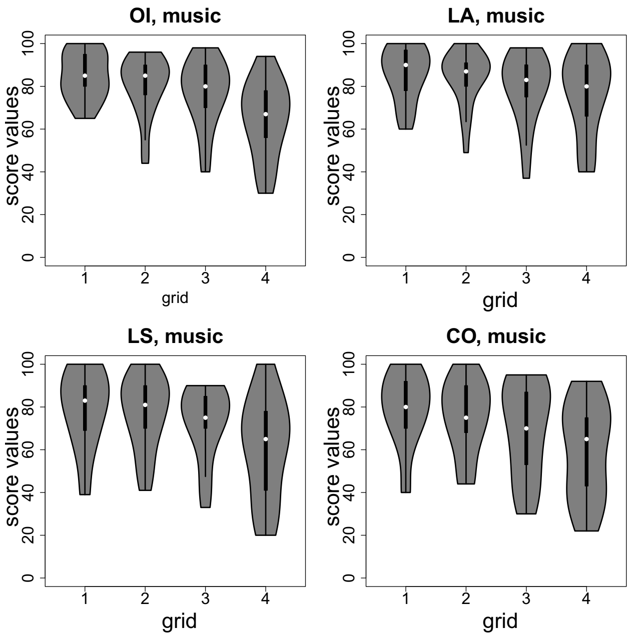

Fig. 8: Ratings of the quality features for the different grids and for the *music signal* as violin plots including boxplots; OI... Overall Impression, LA... Localization Ability, LS... Localization Stability, CO... Coloration.

authenticity would most likely lead to a much more critical evaluation [3].

5. Conclusion

The evaluation with regard to the detection of abnormalities and on the quality assessment shows that the test persons are

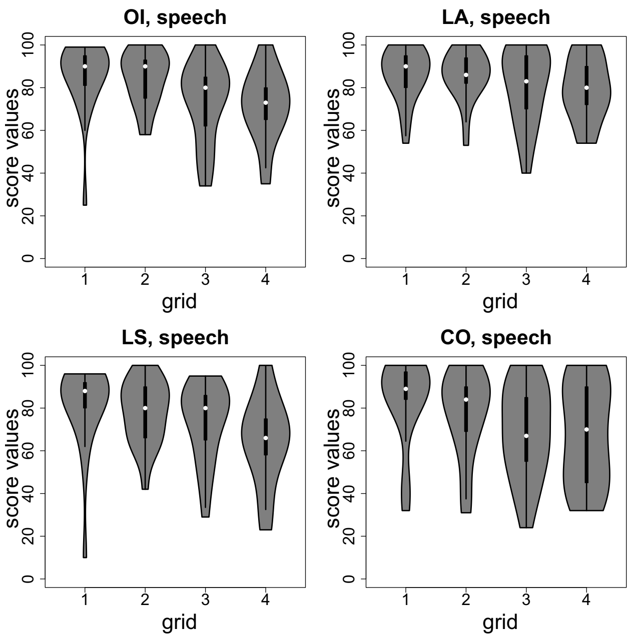

Fig. 9: Ratings of the quality features for the different grids and for the *speech signal* as violin plots including boxplots; OI... Overall Impression, LA... Localization Ability, LS... Localization Stability, CO... Coloration.

quite capable of detecting abnormalities in the various grid conditions. The low-resolution grids are rated worse than the high-resolution grids. However, this effect is much less pronounced for the individual quality features and for the overall impression. It seems to be the case that although the test persons recognize abnormalities, these only lead to a slight decrease of the perceived quality and plausibility of the scene. A further evaluation of authenticity and plausibility can provide information on this. This will also include more critical test signals (e.g. noise), a comparison between uniform grids and quasi-continuous provision of BRIRs. This will be done at a later date.

For a further evaluation in this field, it is intended to consider the assessments for the individual test persons. It can already be seen that the test persons are not homogeneous in their assessments. There are people who generally rate more critically than others.

Nevertheless, the results show that the use of non-uniform grids leads to a suitable position-dynamic binaural synthesis. The number of BRIRs used can be significantly reduced while maintaining a plausible spatial auditory perception. The presented method of non-uniform grids can be applied to any local provision of BRIRs. Furthermore, the method can be used to determine perception differences and perception thresholds in relation to directional and distance perception with a moving listener.

6. Acknowledgment

For their participation in the listening tests and for their interest in research, we would like to thank the test persons. Furthermore, we thank the students of the lecture "Advanced Psychoacoustic" of the Media Technology course at the TU Ilmenau. Special thanks go to Georg Götz and Samaneh Kamandi for working on their media project [4] in our group.

7. References

[1] Daniel H. Ashmead, Deford Leroy, and Richard D. Odom. 1990. Perception of the relative distances of nearby sound sources. *Perception and Psychophysics* 4, 47 (1990), 326–331.

[2] Karlheinz Brandenburg, Estefania Cano, Florian Klein, Thomas Koellmer, Hanna Lukashevich, Annika Neidhardt, Ulrike Sloma, and Stephan Werner. 2018. Plausible Augmentation of Auditory Scenes using Dynamic Binaural Synthesis for Personalized Auditory Realities. In *Audio Engineering Society Conference on Audio for Virtual and Augmented Reality, Redmond, USA*.

[3] Fabian Brinkmann, Alexander Lindau, and Stefan Weinzierl. 2017. On the authenticity of individual dynamic binaural synthesis. *J. Acoust. Soc. Am.* 142, 4 (oct 2017), 1784–1795. https://doi.org/10.1121/1.5005606

[4] Georg Goetz and Samaneh Kamandi. 2018. Optimization of the number and spatial distribution of binaural room impulse responses in an augmented auditory reality application. Media Project, Technische Universitaet Ilmenau, Electronic Media Technology Group.

[5] William M. Hartmann and Andrew T. Wittenberg. 1996. On the externalization of sound images. *J. Acoust. Soc. Am.* 6, 99 (1996), 3678–3688. https://doi.org/10.1121/1.414965

[6] Gavin Kearney, Claire Masterson, Stephen Adams, and Frank Boland. 2009. Towards Efficient Binaural Room Impulse Response Synthesis. In *EAA Symposium on Auralization, Espoo, Finland*.

[7] A. W. Mills. 1958. On the minimum audible angle. *J. Acoust. Soc. Am.* 4, 30 (1958), 237–246. https://doi.org/10.1121/1.1909553

[8] Christina Mittag, Stephan Werner, and Florian Klein. 2017. Development and Evaluation of Methods for the Synthesis of Binaural Room Impulse Responses based on Spatially Sparse Measurements in Real Rooms. In *43. Jahrestagung fuer Akustik, DAGA, Kiel, Germany*.

[9] Brian C. J. Moore. 2012. *An introduction to the psychology of hearing*. 6th edition, Emerald Group Publishing Ltd, London, United Kingdom.

[10] Georg Plenge. 1972. Ueber das Problem der Im-Kopf-Lokalisation. *Acustica* 26, 5 (1972), 241–252.

[11] Christoph Poerschmann, P. Stade, and J.M. Arend. 2017. Binauralization of omnidirectional room impulse responses-algorithm and technical evaluation. In *20th Int. Conf. on Digital Audio Effects (DAFx), UK*. 345–352.

[12] Lauri Savioja, Jyri Huopaniemi, Tapio Lokki, and Riitta Vaeaenaenen. 1999. Creating Interactive Virtual Acoustic Environment. *J. Audio Eng. Soc.* 47, 9 (1999), 675–705.

[13] Zora Schaerer and Alxander Lindau. 2009. Evaluation of Equalization Methods for Binaural Signals. In *Audio Engineering Society Convention 126*. http://www.aes.org/e-lib/browse.cfm?elib=14917

[14] W. E. Simpson and D. Stanton Lee. 1973. Head movement does not facilitate perception of the distance of a source of sound. *The American Journal of Psychology* 1, 86 (1973), 151–159.

[15] Simone Spagnol, Rebekka Hoffmann, Arni Kristjansson, and Federico Avanzini. 2017. Effects of stimulus order on auditory distance discrimination of virtual nearby sound sources. *J. Acoust. Soc. Am.* 4, 141 (2017), 375–380.

[16] Stephan Werner. 2018. *Ueber den Einfluss kontextabhaengiger Qualitaetsparameter auf die Wahrnehmung*

von Externalitaet und Hoerereignisort [On the influence of context-dependent quality parameters on the perception of externality and auditory event location]. Ph.D. Dissertation. Technische Universitaet Ilmenau, Faculty of Electrical Engineering and Information Technology, urn:nbn:de:gbv:ilm1-2018000672, Ilmenau, Germany.

[17] Stephan Werner, Florian Klein, Thomas Mayenfels, and Karlheinz Brandenburg. 2016. A Summary on Acoustic Room Divergence and its Effect on Externalization of Auditory Events. In *8th International Conference on Quality of Multimedia Experience (QoMEX), Lisbon, Portugal*. https://doi.org/10.1109/QoMEX.2016.7498973

[18] Stephan Werner, Annika Neidhardt, Florian Klein, and Karlheinz Brandenburg. 2018. Comparison of Different Methods to Create an Interactive Augmented Auditory Reality Scenario Using Sparse Binaural Room Impulse Response Measurements. In *44. Jahrestagung fuer Akustik, DAGA, Garching, Germany*.

[19] Stephan Werner, Mina Voigt, Florian Klein, and Annika Neidhardt. 2018. A position-dynamic binaural synthesis of a multi-channel loudspeaker setup as an example of an auditory augmented reality application. In *30th Tonmeistertagung VDT International Convention, Cologne, Germany*.

[20] Andreas Zimmermann and Andreas Lorenz. 2008. LISTEN: a user-adaptive audio-augmented museum guide. *User Modeling and User-Adapted Interaction* 18, 5 (2008), 389–416. https://doi.org/10.1007/s11257-008-9049-x