DOI: 10.61782/fa.2025.0500

FORUM ACUSTICUM EURONOISE 2025

INVESTIGATING THE EFFECTS OF SOUND SOURCE OCCLUSION IN A REAL DYNAMIC SCENARIO

Tatiana Surdu[1*] **Lukas Treybig**[1] **Florian Klein**[1,3]
Rakesh Rao Ramachandra Rao[2] **Stephan Werner**[1]
[1] Electronic Media Technology Group, Technische Universität Ilmenau, 98693 Ilmenau, Germany
[2] Audiovisual Technology Group, Technische Universität Ilmenau, 98693 Ilmenau, Germany
[3] Fraunhofer Institute for Digital Media Technology IDMT, 98693 Ilmenau, Germany

ABSTRACT

In the field of auditory mixed reality applications, achieving plausible and immersive virtualized scenes is the primary goal. Progressively, the complexity of the considered scenes is increasing, including, for example, the occlusion of sound sources. Thus, perceptually motivated rendering strategies that cover such phenomena are needed. Hence, this work investigates the perceived transitions between occluded and non-occluded regions around an obstacle in a dynamic scenario, considering distance dependencies and direction of movement. A listening test was conducted, where subjects were required to indicate the transition between occlusion and non-occlusion during walking in a real environment. The test was conducted semi-blind in an acoustically dry room, considering different distances for the source and receiver. For each test condition, the subjects walked along a line parallel to the wall, identifying the area or point of interest. White noise and male speech were used as test samples, played back over two real loudspeakers. The results indicate that changes induced by the presence of an obstacle are distance-, walking-direction, and stimuli-dependent. Further, the perceived transition between different acoustic regions showcases similarities to boundaries proposed by standard geometric models.

Keywords: perception, occlusion, mixed reality

***Corresponding author**: *tatiana.surdu@tu-ilmenau.de.*

1. INTRODUCTION AND RELATED WORK

Daily encountered auditory scenarios are characterized by a wide array of complex acoustic phenomena. Active sources emit sound waves that can propagate freely or be reflected or absorbed depending on the acoustic properties of the elements they encounter. In realistic environments, in addition to the walls, other obstacles such as furniture or clutter also affect the properties of the sound. Placing a sound source behind an obstacle induces audible effects for an active receiver even for partial occlusion [1]. As currently developed mixed reality (MR) systems involving a dynamic, free exploration of the scene aim to generate plausible and immersive virtual scenes, accounting for the reproduction of higher complexity acoustic phenomena that match the scene becomes highly important. For an accurate and exact scene rendering, a high computational cost is to be expected, which is especially critical if considering real-time applications.
Thus, finding perceptually driven approaches to develop simpler rendering solutions becomes important. These optimizations should take into account the spatial distribution (e.g. position and distance) of the active configuration elements such as the source, obstacle, and receiver under specific acoustic settings. The combination of the factors (source, obstacle, receiver, and environment) and their corresponding properties can lead to the sound waves being attenuated, reflected, transmitted, or diffracted around the occluding element. Diffraction is characterized by the formation of three different regions around an obstacle edge based on the signal components that reach them [2,3]. Region I contains incident and reflected waves. Region II is characterized by direct sound and Region III also known as the 'Shadow Zone' contains no reflected or inci-

dent sound components. All regions are characterized by diffraction components [2, 3]. This presents an ideal representation. When analyzing the wave behavior, Fresnel zones, which describe the interference output of the different signal components, indicate a more gradual character of the regions and the transition between them. This formation can be observed for loudness changes in respect to noise barriers [4]. Further, this could also be translated to coloration changes based on signal spectral characteristics. Hence, an adaptation of the ideal representation of regions is required to capture such potential effects.

From a perceptual perspective, the presence of an obstacle can induce audible effects targeting the signal coloration or feeling of spaciousness [5]. Comb-filter effects [6] are also known to emerge, however, their rendering can be neglected for dynamic scenes [7]. Perceived source position shifts, as a result of the formation of new sources along the edge, have also been recorded under anechoic conditions [6, 8]. A decrease in perceived loudness along an obstacle was also recorded by Rungta et al. [9] in a simulated scenario. To virtualize auditory scenes in a plausible manner, such perceptual cues need to be accounted for. Several methods are proposed to achieve this goal. Mannall et al. [7] performed a perceptual comparison of different diffraction rendering solutions considering both static and dynamic scenes. Their results indicate that rendering solutions do achieve different degrees of naturalness, but adaptive solutions are required based on position or degree of motion. The dependency between diffraction rendering and the active considered environment was also addressed by the authors in a later study considering virtualized scenes [10]. Their work unveiled perceived differences based on the reverberation conditions.

In addition to the spatial distribution of the active configuration elements, the direction of walking from/to the occluder can also have an impact on the perception of the transition between different acoustic regions. This aspect is supported by results from Yabe et al. [11], which report that a memory mechanism that integrates successive auditory input into auditory event percepts has a temporal window of 160-170 ms.

While different solutions are proposed for simulating the induced effects of obstacles, much remains unknown regarding the additional effects of more reverberant environments or the diffraction region formation and corresponding effects, under such settings. This becomes particularly critical for augmented reality scenarios, which include both real and virtual elements, thus providing active references for a user. In the present work, the effect of source and receiver distance relative to an obstacle and the influence of the direction of lister movement into the occluded area or out of the occluded area are investigated using a real setup in a non-anechoic environment. The considered space is a dry acoustically treated laboratory. The introduced effects are assessed by using both attributes that quantify and characterize the phenomena, as well as a clicking paradigm that indicates the perceived transition position between different 'regions'.

2. EXPERIMENTAL SETUP

To assess the perceived effects induced by an obstacle, two listening tests were conducted. The tests took place in a dry, non-anechoic, acoustically controlled environment with a reverberation time of $0.21\,\mathrm{s}$. Its design follows the recommendations proposed in ITU-R BS.1116-3 [12]. The room has the following dimensions: $8.4\,\mathrm{m} \times 7.6\,\mathrm{m} \times 2.8\,\mathrm{m}$, resulting in a total volume of $178.75\,\mathrm{m}^3$.

To investigate the effects induced by occlusion, an artificial reflective wall consisting of painted particle boards joined together by aluminum bindings was chosen as an obstacle. The wall had a length of approximately $3\,\mathrm{m}$, a height of $2.5\,\mathrm{m}$, and a width of $0.015\,\mathrm{m}$. The construction was built against the laboratory wall to avoid the lateral bending of waves. As an active real sound source, the ME Geithain Studio RL-906 loudspeakers were selected, and connected to the computer using the MOTU M8 interface. The sources were placed at different distances away from the wall, oriented towards the lateral edge, to enable the playback throughout the test without moving the loudspeakers. An overview of the loudspeaker placement can be seen in Tab. 1a.

To capture the tracking data continuously throughout the test, the Qualisys Motion Capture tracking system was used. The system was operating at a $150\,\mathrm{Hz}$ sampling frequency and was connected to the test script using an openly available Python SDK [1]. To ensure a stable mount of the tracking body the headphone model BK211 [13] was used. This was also done to distract the subjects from the real reproduction form, considering a minimal perceptual distortion [14]. The setup was concealed from the subjects using a chiffon curtain, which has a low impact over the transmission of sound [15]. To further ensure the removal of visual cues, the lights in the laboratory were dimmed as chiffon is a sheer material. In front of the wall

[1] `https://github.com/qualisys/qualisys_python_sdk`

several receiver walking paths were marked, considering different distances and lengths as presented in Tab. 1b. An overview of the entire test setup under normal light conditions, can be visualized in Fig. 1.

Table 1: Overview of the different test layout parameters: (a) The source distances and orientation relative to the wall. φ_S symbolizes the angle between the loudspeaker and the wall edge. (b) The chosen walking paths with their corresponding distances relative to the wall and length. Lines 1-4 were used in the large-scale test and Lines A and B in the expert test.

(a)

Source	Distance	φ_S
S1	2 m	53°
S2	3.5 m	68°
SA	1.5 m	45°
SB	3 m	65°

(b)

Path	Distance	Length
Line 1	1.5 m	3 m
Line 2	2 m	3 m
Line 3	2.5 m	3 m
Line 4	3 m	3 m
Line A	1.75 m	4 m
Line B	2.75 m	4 m

Figure 1: Overview of the test setup under normal light conditions. The green markings indicate Lines 1 to 4 used in the 'Large-scale Test' and the pink markings indicate Line A and B used in the expert test. Minimal lighting was used during the tests.

3. LISTENING TEST METHODOLOGY

The goal of the listening test was to investigate where subjects perceive possible changes induced by an obstacle, and if present, to characterize them using acoustic features. We conducted two listening tests for this investigation. For both tests, the task of the subjects was to walk along the designed paths as indicated in Tab. 1b and click using a pointer when a point or area of change between acoustic regions along the paths were identified. To optimize the testing methodology, a pre-test was conducted, including expert listeners only. The findings and redesign strategies are presented in the following section.

3.1 Expert Test

For this test, Line A and Line B were selected as indicated in Tab. 1b. Two sound sources were utilised, namely SA and SB as described in Tab. 1a. As test stimuli, two samples were used: female speech and pink noise bursts. Overall, 8 trials were designed, and presented in a randomized manner without repetition. After each test condition was completed, the subjects rated the perceived change between the 'zones' using SAQI [16], excluding the 'Artifacts' section.
Six subjects with ages ranging from 28 to 38 years (mean = 32.83), without any self-reported hearing impairments, participated. The subjects included research staff working in the field of virtual acoustics, focusing on the development of binaural rendering systems and their perceptual evaluation, with work experience ranging from 3 to 15 years. To further assess the level of expertise, the subjects were required to submit ratings related to the frequency of listening test conception, test participation, and general listening expertise using a 5-point scale. The average scores are given as follows: 3.83, 5 and 4.3. From the 6 subjects, 5 are also musically active. One of the goals of the expert listening test, was to identify the most important rating attributes from SAQI for the present acoustic scenario. To subsample the questionnaire, the number of occurrences in the rating was counted. Considering 6 subjects and 8 trials, the maximum count is 48. Seven attributes with the highest count were considered to be used for the large-scale test. These include: Difference (48), High-frequency tone colour (46), Tone colour bright-dark (45), Loudness (35), Mid-frequency tone color (24), Sharpness (24), and Localizibility (24). The results are in line with previous works, which recorded perceived changes in the signal spectral properties [5] and loudness [9]. Based on the seven attributes, four main features were sub-sampled and redefined for the large-scale test: 'General Difference', 'Loudness', and 'Localizibility', and a custom attribute 'Spectral Changes' (includes the previously addressed frequency-related attributes). Reducing the number of attributes was needed to not exceed the testing duration for the rating of more conditions. Since listening tests should consider a larger mixed participant pool, reducing the terminology complexity could be bene-

ficial for naive subjects, which additionally motivated the creation of the custom attribute. Furthermore, for the majority of attributes rated, which use a bipolar scale, only one side of the scale was utilized in the expert test. This motivated the usage of a unipolar 5-point scale for the final custom questionnaire. It is important to note that the reported perceived position shifts were not related to the presence of the obstacle but to a reflective projector, which induced changes in the perceived elevation. This aspect motivated adapting the walking path length from $4\,\mathrm{m}$ to $3\,\mathrm{m}$, stopping before the projector location. Other layout-related changes included the usage of more walking paths to ensure more granularity in the data collection, as well as different distances for both source and receiver.

3.2 Large-scale Test

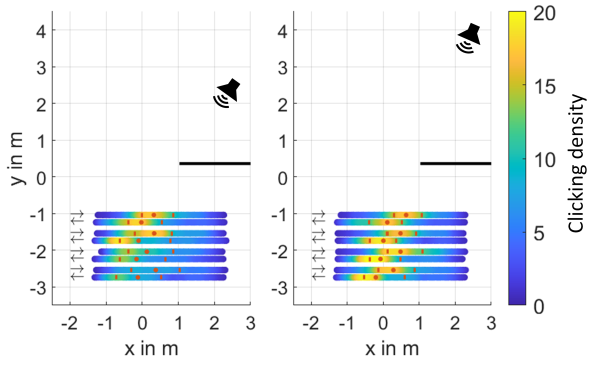
Using the modifications derived from the expert test, a new large-scale test was designed and conducted. For this test, 4 walking paths: Lines 1 to 4, were included as described in Tab. 1b. Different source distances as indicated for S1 and S2 in Tab. 1a were considered. To investigate the effect of the signal spectral properties, male speech, and white noise were used as test stimuli. 32 trials were designed, considering the parameter variations as mentioned before, including a repetition of each condition. The trials were presented randomly, with each participant receiving a different presentation order.
Prior to the test beginning, the subjects underwent a training session. In the initial part, the attributes were introduced as defined in the SAQI questionnaire. For the 'Spectral changes' attribute, a cumulation of definitions were presented related to frequency-related aspects. In the next step, the test paradigm was introduced. SB was utilized as a training source, playing back a saxophone sample. The subjects were instructed to walk along Lines A and B and familiarize themselves with the different 'zones', corresponding transition phenomena, and the clicking rating paradigm.
For each trial, the subjects walked along the designated path at least once. As opposed to the expert test, a rating was provided for each direction of walking, at least once. If no change was perceived while walking along the line, the subjects provided a 'blank' rating outside the paths. Subsequently, the subjects additionally used the four custom items to evaluate the specific condition.

4. TEST RESULTS

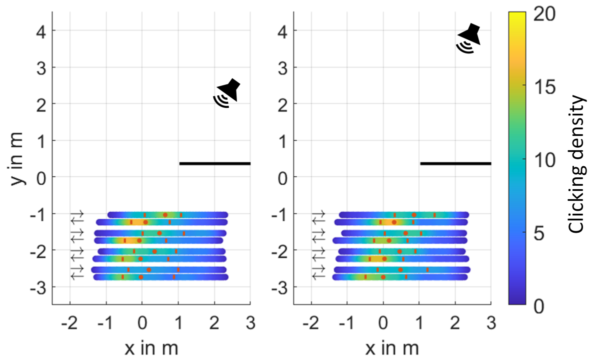
For the large-scale test, 30 subjects (6 F and 24 M) with ages ranging from 20 to 38 years (mean = 28.8) participated in the study. The subjects did not report any hearing impairment except for 2 participants who suffer from tinnitus. The subject pool consisted of research staff as well as students from the university, thus including naive and expert participants. Participants who were not employed at the university were monetarily compensated. The expertise was assessed based on experience working in the field of acoustics, and the subjective evaluation of MR systems, as well as, music activity (playing or producing), resulting in 8 subjects representing the 'expert group'. The clicking and tracking data for one subject data was removed from the first part of the analysis as technical issues were discovered after the test.

4.1 Detection of occlusion

Fig. 2 illustrates the distribution of ratings based on the signal type, sources, and lines, categorized by the direction of movement. The direction of movement is represented by arrows and was determined using the average speed calculated from 100 samples prior to the submitted evaluation. The scaling of the distribution of evaluation frequencies is shown by the color bar. In addition, the quartiles of the distributions and the medians are marked in red. Click data given outside the paths, indicating no perceived change were not considered for the density calculation. The clicking frequency and distribution showcase the perceived transition between acoustic zones. A strong dependency can be observed between the rating and the distance relative to the wall for source and receiver. For S1, placed closer to the wall, the density of ratings decreases for Lines 3 and 4. A possible explanation for this can be provided by the geometric model representation [2, 3] of the three regions forming around the edge of an obstacle as described in Sec. 1. For S1 ($\varphi_S = 53°$), Lines 3 and 4 are mostly in the 'Shadow' region. Thus, no significant audible effects are noticeable when walking along the more distant paths for S1. For S2, placed further away, both regions II and III, as well as the shadow zone, are included. This explains the high click density even for larger distances. It is important to note that these considerations, when analyzing the regions according to the geometric model, do not take into account the signal spectral characteristics. A clear difference in the rating can be observed, for the two chosen signals. For white noise, the number of clicks is higher than for the speech



(a) white noise



(b) speech

Figure 2: Distributions of the clicking density as a function of the direction of movement for the white noise (a) and the speech signal (b); **Arrows** mark the direction of movement; **Red**: median and quartiles.

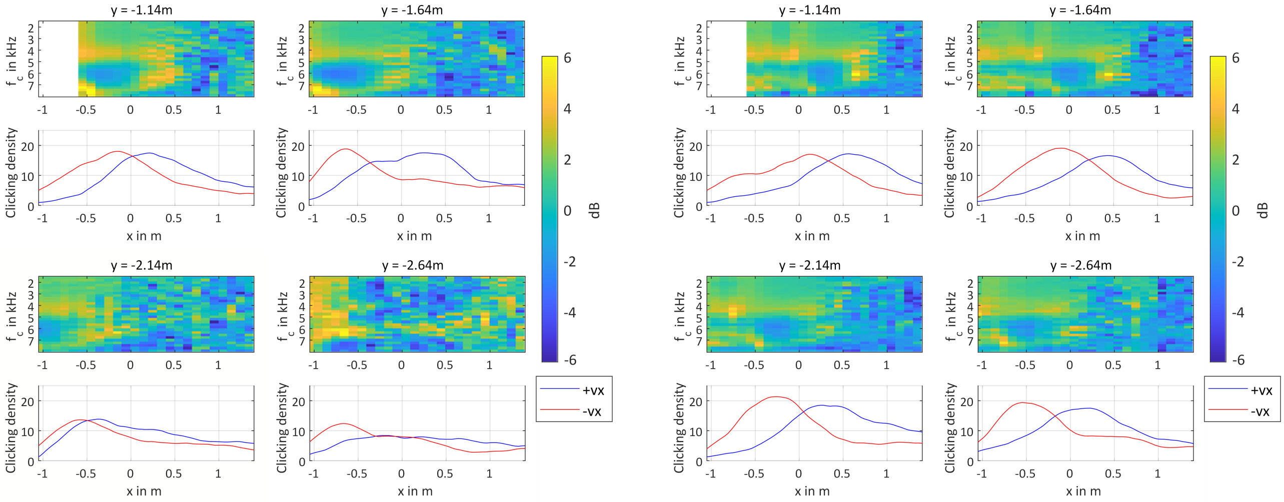
across all conditions. This can be explained by the spectral characteristics of each sample and how the presence of an obstacle has a different effect based on the considered frequency. In addition, the perceived effects strongly depend on the direction of walking. The results indicate a shift in the perceived area of change, independently of stimulus or source and receiver distance relative to the wall. The shift in the click density as a function of the direction of movement could be related to the direction-dependent transmission properties of the outer ear. To investigate this, spatial room impulse responses (SRIRs) measurements using the platform described in [17] are utilized. Using this data, binaural room impulse responses (BRIRs) were synthesized using the Spatial Decomposition Method (SDM) [18]. The BRIRs included the orientation in the x-direction for 0° and 180°. To account for perceptual variation considering the direction of walking, the spectral difference (S) between the two ears, directed at the loudspeaker for the two orientations, was calculated following Eqn. (1).

$$\Delta S(f_c, x) = S_{0,\text{left}}(f_c, x) - S_{180,\text{right}}(f_c, x) \quad (1)$$

The spectral differences (S) for the two sound sources S1 and S2 as well as for the different distances are shown in the upper parts of Fig. 3. A specific pattern of reduced and increased differences of approx. -/+ 4 dB in the range of approx. 5 kHz to 7 kHz for neighboring areas of the walking lines is visible. For increasing distances from the wall, this pattern shifts towards the non-occluded region (to the left on the x-axis). For the nearby sound source S1, the pattern shifts out of the measured range for further away distances. The lower parts of Fig. 3 show the evaluation behavior of the test persons for the indication of a perceived difference (click density) during movement on the lines for both walking directions. This results in maxima at specific positions on the walking lines. The maxima for the direction +*vx* tend to correlate with the positive spectral difference and the maxima for the direction -*vx* tend to correlate with the negative difference. It is expected that the test persons use different spectral coloration cues to detect a transition between the masked and the direct sound region. The transition point is also not at the same position on the line but depends on the direction of walking. The results appear to be consistent with the model assumption of a gradual transition from the occluded region ('Shadow Zone') to the direct sound region as described in section I. This suggests that when moving from the direct sound region to the occluded region, only a relatively strong occlusion becomes perceptible. On the other hand, when moving from the occluded region to the direct sound region, non-occlusion is only detected close to the direct sound region.

4.2 Characterization of occlusion effects

To investigate what factors aided the participants in perceiving the transition between different acoustic zones (occluded and non-occluded), the participants were asked to rate the perceived degree of change of four different attributes that were subsampled based on the results of the expert test. These include "General Difference", "Spectral Changes", "Localizibility", and "Loudness". All the attributes were rated on a 5-point scale where a rating of 1 indicates 'No Change' and a rating of 5 indicates 'Large Change'. The distribution of the attributes across different



(a) S1 (b) S2

Figure 3: **Upper parts**: Difference of the BRIR spectra of the ear facing the loudspeaker with respect to the orientation 0° and 180° along the x-axis for the individual lines according to Eqn. (1); **Lower parts**: Distributions of the click density as a function of the direction of movement for white noise, for the sound sources S1 (a) and S2 (b).

distances, sources, and stimuli can be found in the supplementary material here [2]. We performed statistical significance testing to investigate if there was a significant impact of any of the independent factors, namley, distance, source, and stimuli, on the perception of degree of change on the four attributes. As the subjective data was not normally distributed, we performed the Scheirer–Ray–Hare test to examine if the attribute ratings were affected by the three independent factors. Scheirer–Ray–Hare test is a non-parametric statistical test that can be applied to data which is not normally distributed and is an extension of the Kruskal-Wallis test. The results of this analysis showed that the effect of each of the three independent factors were significant across the ratings of all attributes. However, all interaction effects were found to be nonsignificant. This indicates that the influence of one independent factor is not dependent on another factor.

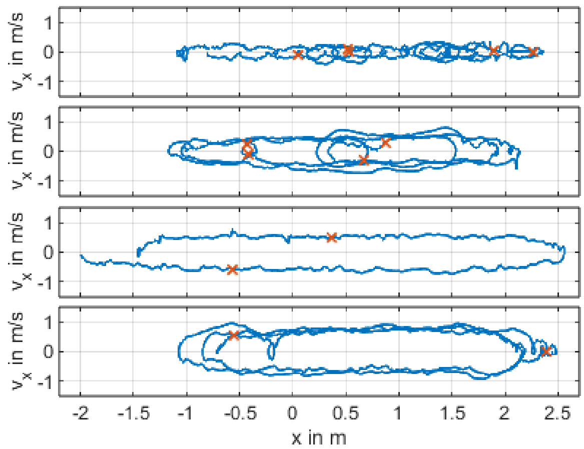


Figure 4: Movement behavior of 4 participants as an example for white noise, with S2 along Line2; **red**: ratings submitted.

5. DISCUSSION

The present work showcases the perceived effects induced by an obstacle, considering different source and receiver distances, walking directions, and test stimuli in a non-anechoic environment. The noise sample, is characterized by a constant intensity across the entire spectrum. As diffraction, absorption, or reflection affect the signal spectra differently, induced changes are more easily audi-

[2] `https://github.com/tuil-emt/SoundSourceOcclusionFA25`

ble for a broadband critical signal. The male speech sample, however, has energy concentrated mostly in the mid-frequency range, with power from 100 Hz until around 1000 Hz. As diffraction is particularly critical for low frequencies [5] and high-frequency components are more attenuated, it becomes more difficult to assess minute changes for the male speech sample.

A dependency on distance, and perception for both receiver and source can be observed in the present results. This is reflected in the rating (click) density presented in Fig. 2. The rating density indicates that the differences between the acoustic regions are more perceivable if the transition between occluded and non-occluded regions, corresponds to the geometric boundary representation (between Region II and Region III) as represented in [2, 3]. For source S1, most of the walking path lengths for lines L3 and L4 fall within the shadow zone and have a lower rating density as no transition could be rated. For source S2, all the lines fall between Region II and Region III with varying perceivable differences recorded with increasing distance. One possible explanation for this could be that the effects, such as the perceived degree of changes for the different attributes, which aid in identifying the transition in acoustic zones, vary significantly across different distances, possibly making it harder for participants to perceive such transitions. This, along with the impact of the stimuli on the perception of the transition of different acoustic zones raises the question of at what distances do the effects of occlusion become negligible. This knowledge is important as it can be beneficial for developing efficient and plausible rendering solutions. In addition to this, from the analysis of the tracking data, a perceptual shift of the boundary between acoustic zones was observed for the investigated test case for the different directions of walking. A possible explanation for this can be the spectral differences due to the head orientation and the amount of information reaching the listener as indicated in Fig. 3. This is in line with the findings of Yabe et al. [11] which suggests that an auditory event percept is formed by integrating successive auditory input within a temporal window of 160-170 ms.

Furthermore, the analysis of the tracking data uncovered differences in subject walking behavior. As Fig. 4 showcases, some subjects walk slower (first subfigure) and focus more on specific areas. Other subjects, on the other hand, (third subfigure from the top) explored the path only once and much faster. This indicates that aspects, such as subject reliability have to be investigated, as subjects not exploring the scene may not provide reliable responses. To alleviate this, using head rotation tracking data for a more robust assessment of the level of effort and attention allocated during the task can be considered.

6. SUMMARY AND FUTURE WORK

In the present work, four factors and their impact on the perception of an obstacle in a non-anechoic environment were investigated. An initial test with expert listeners was designed as a means of optimizing the test design for a second, larger-scale test using a mixed participant pool. The results from the large-scale study indicate a dependency between the test stimuli and how subjects perceive the transition between different acoustic regions forming around an obstacle based on the direction of walking. Considering white noise, subjects can more easily perceive acoustic changes. In addition, the relative spatial placement of the source and walking paths plays an important role. The results indicate that geometric boundaries between different regions can reliably mark where the shadow zone appears. However, the direction of walking indicates a shift in the boundary position, which could be an indicator for multiple perceivable boundaries between different regions. For increasing distances, subjects still perceive the effects between different zones, which motivates the need to establish audibility thresholds with respect to the distance from the obstacle in future studies. As the end goal is the development of a perceptually optimized MR system, rendering simplifications should be considered. This is particularly challenging for 6 DoF dynamic representations, considering the possibility of real present references in the scene. Binaural measurements using a dummy head could be employed in real environments with a varying degree of acoustic complexity to facilitate the perceptual validation of other rendering methods which focus on computational simplifications. For more complex scenarios, the impact of source position relative to the wall could be further investigated. Similarly, obstacles with varying properties such as shape, size, or material could be investigated to asses the perceptual impact for MR systems. The present work investigated the effects in a relatively dry environment, recording perceived differences. However, the degree of reverberation and its impact on perception can be further investigated as suggested by other studies [5, 10]. As the present work unveiled differences in the rating and exploration behaviour for the subjects future studies could include additional measures that capture the response certainty.

7. ACKNOWLEDGMENTS

The authors would like to thank the technical staff for their help in preparing the test setup, the subjects who participated in the test and Sebastian Schwarze for co-supervising the listening tests. The research staff is funded by the Projects: *"Co-Humanics"* and *"Multiparties"*.

8. REFERENCES

[1] M. K. Russell and S. Brown, "Using sound to create and detect occlusion of an unseen sound source," *Auditory Perception & Cognition*, vol. 2, no. 4, pp. 207–229, 2019.

[2] R. R. Torres, U. P. Svensson, and M. Kleiner, "Computation of edge diffraction for more accurate room acoustics auralization," *The Journal of the Acoustical Society of America*, vol. 109, no. 2, pp. 600–610, 2001.

[3] P. Nikolaou and P. Menounou, "Analytical and numerical methods for efficient calculation of edge diffraction by an arbitrary incident signal," *The Journal of the Acoustical Society of America*, vol. 146, no. 5, pp. 3577–3589, 2019.

[4] A. M. Lacasta, A. Peñaranda, and I. R. Cantalapiedra, "Green streets for noise reduction," in *Nature based strategies for urban and building sustainability*, pp. 181–190, Elsevier, 2018.

[5] R. Torres, N. de Rycker, and M. Kleiner, "Edge diffraction and surface scattering in concert halls: physical and perceptual aspects," *Journal of Temporal Design in Architecture and the Environment*, vol. 4, no. 1, pp. 52–58, 2004.

[6] H. Farag, J. Blauert, and O. Abdel Alim, "Psychoacoustic investigations on sound-source occlusion," in *Audio Engineering Society Convention 114*, Audio Engineering Society, 2003.

[7] J. Mannall, O. Das, P. Calamia, and E. De Sena, "Perceptual evaluation of low-complexity diffraction models from a single edge," in *AES Conf.: 2022 AES Int. Conf. on Audio for Virtual and Augmented Reality*, Audio Engineering Society, 2022.

[8] C. Kirsch and S. D. Ewert, "Effects of measured and simulated diffraction from a plate on sound source localization," *The Journal of the Acoustical Society of America*, vol. 155, no. 5, pp. 3118–3131, 2024.

[9] A. Rungta, S. Rust, N. Morales, R. Klatzky, M. Lin, and D. Manocha, "Psychoacoustic characterization of propagation effects in virtual environments," *ACM Transactions on Applied Perception (TAP)*, vol. 13, no. 4, pp. 1–18, 2016.

[10] J. J. Mannall, P. Calamia, L. Savioja, A. Neidhardt, R. D. Mason, and E. De Sena, "Assessing diffraction perception under reverberant conditions in virtual reality," in *International Conference on Audio for Virtual and Augmented Reality (AVAR)*, 2024.

[11] H. Yabe, M. Tervaniemi, J. Sinkkonen, M. Huotilainen, R. J. Ilmoniemi, and R. Näätänen, "Temporal window of integration of auditory information in the human brain," *Psychophysiology*, vol. 35, no. 5, pp. 615–619, 1998.

[12] International Telecommunication Union, "Methods for the subjective assessment of small impairments in audio systems," Recommendation ITU-R BS.1116-3, ITU, 2015.

[13] V. Erbes, F. Schultz, A. Lindau, and S. Weinzierl, "An extraaural headphone system for optimized binaural reproduction," *Fortschritte der Akustik - DAGA 2012*, pp. 313–314, 2012.

[14] C. Schneiderwind, A. Neidhardt, and D. Meyer, "Comparing the effect of different open headphone models on the perception of a real sound source," in *AES Conf.: 2020 AES Int. Conf. on Audio for Virtual and Augmented Reality*, Audio Engineering Society, 2020.

[15] J. Heeren, G. Grimm, S. Ewert, and V. Hohmann, "Video screens for hearing research: Transmittance and reflectance of professional and other fabrics," *arXiv preprint arXiv:2309.11430*, 2023.

[16] A. Lindau, V. Erbes, S. Lepa, H.-J. Maempel, F. Brinkman, and S. Weinzierl, "A spatial audio quality inventory (saqi)," *Acta Acustica united with Acustica*, vol. 100, no. 5, pp. 984–994, 2014.

[17] G. Stolz, S. Werner, F. Klein, L. Treybig, A. Bley, and C. Martin, "Autonomous robotic platform to measure spatial room impulse responses," *Fortschritte der Akustik - DAGA 2023*, pp. 59–61, 2023.

[18] S. V. A. Gari, J. Arend, P. Calamia, and P. Robinson, "Optimizations of the spatial decomposition method for binaural reproduction," *Journal of the Audio Engineering Society*, vol. 68, no. 12, pp. 959 – 976, 2020.