Proceedings of the ASME 2022
International Mechanical Engineering Congress and Exposition
IMECE2022
October 30-November 3, 2022, Columbus, OH

IMECE2022-97044

PRESERVING AUDITORY CUES FOR HUMAN ECHOLOCATION TRAINING: A GEOMETRICAL ACOUSTICS STUDY USING A BENCHMARK DATASET (BRAS)

Jonas Karlberg[1,†,*], Alessia Milo[2,†], Finnur Pind[2], Runar Unthorsson[1]

[1]The Faculty of Industrial Engineering, Mechanical Engineering and Computer Science, University of Iceland, Reykjavik, Iceland
[2]Treble Technologies, Reykjavik, Iceland

ABSTRACT

Human echolocation is a method mainly used within the blind community to navigate using sound emission and analysing the returning echoes from the surrounding environment. Echolocation is predominantly trained by orientation and mobility instructors at visual rehabilitation centres. However, systematic guidelines or protocols focusing on the requirements of room acoustic simulations to accurately represent the auditory cues necessary in a virtual training environment for echolocation have not yet been developed.

This paper sets out to investigate the use of geometrical acoustic (GA) calculations for a virtual echolocation training system comparing the measurements from the Benchmark for Room Acoustical Simulation (BRAS) dataset with other GA calculations outcomes. Three simple and one complex test scenes are chosen from the dataset. The calculation settings are optimised for each test scene considering the complexity of the scene, room volume and acoustic phenomena.

The monaural room impulse responses from the simulations are analysed with respect to the timing of the reflections and the level relations between the reflections and the resulting frequency response for each scene. These are subsequently compared with each measured counterpart. The paper discusses the results, their limitations, and provides recommendations on the use of GA calculation tools for echolocation training scopes.

INTRODUCTION

All people with normal hearing possess the ability to use echolocation but, because sighted individuals learn about their environments visually, they often do not readily perceive echoes from nearby objects. Human use of echolocation is normally associated with blind individuals and there are well known examples of blind individuals that have successfully mastered this ability. Most notable are Daniel Kish and Ben Underwood. Daniel Kish and his organisation have taught a form of echolocation to at least 500 blind children around the world. Echolocation is predominantly trained by orientation and mobility instructors at visual rehabilitation centres[1, 2]. Nicole Holmes proposed a guideline protocol for echolocation training describing different test scenarios with increasing complexity[3].

Globally, there are 285 million visually impaired people, whereof 39 million who are blind. A recent survey among people with visual impairment in Norway (N=736) found that 50 percent of the respondents could be considered lonely, where 20 percent had a high degree of loneliness. The incidence of feeling lonely is substantially higher than in the general population, in all age groups. Another significant finding is that many visually impaired people avoid seeking help because they find it shameful not be able to manage on their own. By acquiring the skill of echolocation, blind people can become more self-reliant, self-directed and less dependent on the aid from others. This can substantially increase their confidence, autonomy, sense of dignity and freedom, and make it easier to participate and contribute to the society[4–6].

There is a large difference between how individuals learn echolocation and some are born with much better orientation skills. Hence, some visually impaired people have been able to become highly proficient in using the echolocation technique while there are others that have not[7, 8].

This paper investigates the use of GA algorithms and their ability to preserve the important auditory cues for use in a virtual echolocation training system.

Auditory cues for human echolocation

In this section a brief overview from a selection of research within the topic of auditory cues used in human echolocation

[†]Joint first authors
[*]Corresponding author: jonaskarlberg@hi.is

is presented. For further reading summaries by Kolarik et al. 2014 and Thaler & Goodale 2016 are found in Ref. [8] and [6] respectively. From the summary by Kolarik et al. 2014 the following cues were presented:

1. Loudness and the energy relation between emission and echo.

2. Time separation pitch, also known as repetition pitch, which is the time between emission and echo. Time separation pitch lay in the range of 1 to 30 ms.

3. Change in timbre or pitch, which is the frequency domain equivalent of time separation pitch. The changes in the frequency spectrum occur due to interference from the returning echoes.

4. Time differences between the ears of the receiver. This is especially important at high frequencies. These cues can be helpful for localising objects.

5. Changes in the reverberation pattern of a reverberant room due to objects placed in such room.

For the investigations in this paper cues 1-3 are considered.

Several studies state that, in general, blind individuals are found to be better at echolocation than sighted ones[6]. There are studies that suggest that blind individuals have enhanced spectral processing capabilities compared to sighted individuals which can be one explanation for this[9–13]. In a study by Muchnik et al. 2009 it was reported that blind individuals trained in echolocation were able to detect the separation of two sounds, with a silent gap in between, down to a separation time of 5 ms[14].

The emission from mouth-clicks usually have peak frequencies in the range of 3 to 8 kHz stated in Ref. [6] as reported by [15, 16]. The frequency range of use in echolocation and its important cues are reported to be above 2 kHz[17, 18]. Participants in studies by DeLong et al.[19] and Hausfel et al.[20] reported that the main cues for determining material properties of an object were pitch and timbre changes.

A virtual echolocation system was implemented by [21]. However, the modelling of room acoustics was absent in this project. On the other hand, Picinali et al.[22] conducted a series of studies on spatial reconstruction of achitectural spaces also through a VR interactive application, finding that head tracking and dynamically choosing the navigation speed is crucial in helping participants understand the spatial configuration of the environments and the location of sound sources. In a distance comparison task, they found no significant differences between the real and virtual navigation modalities. Therefore, adding virtual acoustics to a training system could add flexibility, in terms of designing virtual test environments, as well as putting the participant and the sound sources, in this virtual scene, into a familiar context i.e. the perception of sound in a room.

Geometrical acoustics

Ray tracing is a common technique used in modelling of wave propagation where emitted sound from a source is represented as a ray and its propagation following the reflection laws from geometrical optics. The use of computer modelling in room acoustics was introduced by Schroeder et al. in 1962[23] and later implemented by Krokstad et al. in 1968[24, 25]. In stochastic ray tracing modelling a Dirac impulse is emitted from a sound source for every ray, carrying a specific energy. When a ray hits a boundary, such as a wall, it changes its path and loses a portion of its energy while travelling in the new direction. The energy loss is decided by the assigned material of the reflecting surface. The receiver in the simulation can be modelled either as a volume or a surface. A volumetric sphere represents an omnidirectional microphone with no spatial information. In order to capture the spatial information either two points need to be modelled or a set of points, creating a microphone array for spherical harmonic representation of the sound field[25].

The image source method is based on the creation of virtual mirrored sources where the direct sound and incoming reflections to the receiver are added[26]. With this approach phase information is included by adding the complex spherical wave amplitude to the histogram. The image sources are created up to a defined reflection order. For example, using a reflection order of two implies that mirror sources are created for the first order image sources.

The secondary source method has been developed to calculate the late part of the impulse response (*IR*) tail[27]. In this approach, a secondary source is created on each surface where rays from an image source hit the boundary. The secondary source from the surface point can be considered to radiate as a hemisphere[28].

The room acoustic calculation software ODEON uses a hybrid approach for the creation of the IR where an image source method and early scattering method are implemented for the early part of the IR and a ray radiosity method based on secondary sources for the late part. More information regarding these methods can be found in Ref. [28] and ODEON user's manual[29].

In classical ray tracing methods, phase is not part of the calculation. However, the image source method, often used for the early part of the IR calculation, captures this in theory. A method for creating the late part of the IR with phase information is described in Ref. [30, 31].

In short, geometrical acoustics using ray tracing is a simplification of the physical phenomena and several different approaches in GA have been developed. In-depth information of the different approaches can be found in various literature such as Ref. [25, 32].

Benchmark for Room Acoustical Simulation

The Benchmark for Room Acoustical Simulation (*BRAS*) dataset, prepared by Aspöck et al. 2019, consists of several

reference scenes isolating acoustic phenomena such as single and multiple reflections as well as diffraction. In addition to these controlled laboratory cases several existing rooms of different size and complexity are included as well[33, 34].

METHOD

Two GA algorithms were chosen for calculation of room impulse responses (*RIRs*). One is based on classical raytracing as implemented by Kulowski[35] using volumetric receivers. The other is the commercial software ODEON[1] which uses a combination of ray tracing methods for the different parts of the IR. For the early part a combination of the image source method and an early scattering method is implemented. For the late reflections the secondary source method is used together with the ray-radiosity approach. The calculated RIRs were then compared against the ones provided by the BRAS dataset. Since timing and energy relation of the reflections are some of the important factors in the presented auditory cues for use in human echolocation these are the objective measures which were studied. The analysis of the differences is hence focused on the arrival time and energy of the reflections in time domain and the resulting frequency response.

The BRAS dataset acted as the comparison reference for the simulated RIRs. Hence, the following provided data was used in the simulations:

- values for absorption, scattering, air humidity and temperature
- RIRs
- 3D geometry models
- source and receiver descriptions

Three simple scenes, with emphasis on an isolated acoustic phenomena, and a complex scene were selected for this study. The simple scenes were the single and multiple reflection cases, *Scene 2* and *3* (named *RS2* and *RS3* in Ref. [34]), situated in an anechoic environment with a rigid plate acting as the reflector and the the diffraction case, *Scene 5* (named *RS5* in Ref. [34]), situated in a hemi anechoic environment with a rigid wall placed between the direct path of the source and the receiver. These scenes were chosen to study the timing of the incoming rays to the receiver and how the corresponding frequency response is affected. The complex scene *Scene 9* (or *CR2* in Ref. [36]), a small unfurnished seminar room, was chosen to observe the overall trends in time and frequency domain. The source-receiver pairs used for the comparison were *LS02* and *MP02* for *Scene 2* and *9*, *LS01* and *MP01* for *Scene 3* and *5*. The scene models were prepared by importing the SketchUp models into Rhino 7 and there creating a new model with simplifications such as removing the thickness of the plates in *Scene 2* and *3* as well as the plates mounting stands. In *Scene 9* the 3D model was simplified by removing detailed features such as the radiators and the sink. A box shaped room was created for *Scene 2*, *3* and *5*, and assigned with absorption coefficient of 1 to suppress any reflections from these boundaries. The dimensions for these boxes were 10 x 10 x 5 m, 20 x 20 x 5 m and 11 x 6 x 5 m for *Scene 2*, *Scene 3* and *Scene 5* respectively. For the diffraction case, being a hemi anechoic scene, the floor was assigned the material provided in the surface description package. An overview of the used materials for each scene and their absorption and scattering values can be seen in Tab. 3.

In Tab. 1 the calculation settings used in ODEON for the simulated test scenes are shown. Air temperature and humidity settings were set as provided by the dataset. The number of early rays was set using the *Precision* setting in the *Room setup*-dialog with a number of early rays set to 500. The source directivity was modelled as a Genelec 8020b loudspeaker (8020c in BRAS reference) for *Scene 2*, *3* and *5*, and as an omnidirectional source for *Scene 9*.

TABLE 1: CALCULATION SETTINGS FOR ODEON

Transition Order	4
Number of late rays	16000
Number of early scatter rays	500
Oblique Lambert scattering	activated
Angular absorption	disabled
Reflection based scatter	disabled
Screen diffraction	activated in scene 5

In Tab. 2 the simulation settings used with the classical raytracing algorithm are shown. Air absorption per meter coefficients in octave bands at 10% relative humidity steps for 20° *C*, were retrieved from Ref. [37]. Interpolation between values was not applied at this stage. The source directivity was modelled as omnidirectional.

Data processing

Following the previously stated auditory cues used in echolocation an objective evaluation of the time signal was carried out with respect to the timing of the peaks in the IR as well as their level relation. The data processing steps was inspired by the method used by Brinkmann et al. from the publication [38]. All data was processed and analysed in MATLAB using the signal processing toolbox[2] and ITA Toolbox by the Institute of Technical Acoustics (RWTH Aachen University)[39]. The IRs were loaded in MATLAB using the *audioread* and *ita_read* routines. In order to extract a monaural IR in ODEON the *Surround export* option was used with one virtual loudspeaker defined, set at the center of the receiver position. It was found that this type of IR generation in ODEON adds a delay to the signal, which might mean that the speaker is not effectively placed where set by the user. However, when studying the arrival time from the echogram in ODEON these values were correct if compared to the measurements. Hence, the direct sound was aligned[3] to the BRAS measurement for the ODEON IRs using *ita_time_shift*.

[1] ODEON Auditorium version 17.

[2] MATLAB version R2022a (9.12.0.1884302).

[3] Shifted -8.8 ms.

TABLE 2: CALCULATION SETTINGS FOR CLASSICAL RAYTRACING

	Scene 2 & 3	Scene 9
Number of rays	10^6	10^5
Receiver radius	0.3 m	0.3 m
Source Directivity	omni	omni

Scene 2 - Single reflection. The IRs were cut at 35 ms, bandpass-filtered and normalised to the max level of the IR using ITA toolbox. The frequency range of the bandpass-filter was set to 125 − 8000 Hz with a filter order of 10.

For the time analysis the peaks in the IRs were detected using *findpeaks* where the signals were normalised to the absolute max value of the signal. The detection thresholds were set using the *MinPeakProminence* option. The time offset between the first reflection and direct sound (FR-DS) was extracted and analysed with respect to the reference.

For the relative level analysis the IRs were squared and summed ±0.5 ms around each peak. The level difference ΔL was then calculated according to Eq.1.

$$\Delta L = 10 \times log_{10}(L_{d,1}/L_{r,1}) - 10 \times log_{10}(L_{d,2}/L_{r,2}) \quad (1)$$

Where $L_{d,1}$ and $L_{r,1}$ are the summed level of the direct sound and reflection for the reference signal and $L_{d,2}$ and $L_{r,2}$ for the compared signal respectively.

Scene 3 - Multiple reflections. The IRs were cut and bandpass-filtered using the same routines and values as in *Scene 2* with the exception of the cutting time which was set to 200 ms.

Scene 5 - Diffraction. In this scene the simulations were performed in ODEON since the classical raytracing algorithm did not include modelling of diffraction. The IRs were bandpassed as previously explained and cut at 35 ms.

Scene 9 - Small seminar room. The IRs were cut at 1.5 s and bandpass-filtered using the same routines as previous scenes. The scene had air humidity of 41.7 % and temperature 19.5° *C*. In the classical raytracing the air humidity was set to 40 % and the air temperature to 20° *C*. For the evaluation of the reverberation time T_{20} the routine *ita_roomacoustics_parameters* was used.

RESULTS

The results from each test scene are presented in the following sections.

In order to assess the reliability of the results with the classical raytracing method, five simulations with identical settings were launched for each scenario and the results prior to IR generation were also inspected. The simulations were not repeated with ODEON at this stage since ODEON requires the user to change settings prior to launching a new simulation.

Scene 2 - Single reflection

In Fig. 1 the simulated and measured IRs for scene 2, single reflection case, are shown. In Tab. 4 the calculated time and level

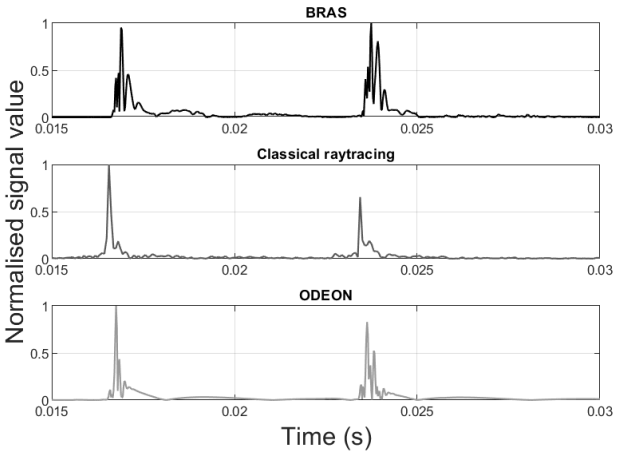

FIGURE 1: THE PROCESSED IMPULSE RESPONSES FROM SCENE 2

difference are shown. FR-DS shows the difference in ms between the first reflection arrival and direct sound arrival. For the BRAS reference, this value is 6.848 ms. The time difference Δt with respect to the reference timing is shown for both the classical raytracing method and ODEON, as well as their percentage with respect to the reference (Δt / FR-DS).

The arrival times are close to the measured reference with a deviation of less than a tenth of a millisecond. For the classical raytracing the small deviation could be explained by the difference in air humidity settings (47 % relative humidity in ODEON and BRAS versus 50 % in classical raytracing). In this test scene, the level of the direct sound in the reference is slightly lower than the level of the reflection. This is due to the directivity of the loudspeaker, oriented towards the plate. In the classical raytracing simulation using an omni source the peak of direct sound is increased due to the extra contribution from the radiation properties. This is seen in the level difference and can be an explanation for the difference between the algorithms performance[4]. Another probable cause for the difference in the results is the different scattering implementations in the algorithms, where the classical raytracing uses random scattering, thereby not achieving as many receiver hits as with the vector based scattering method in ODEON.

With respect to the classical raytracing simulation robustness, it was found that the standard deviation between the levels at all frequency bands across five simulations was less than 0.5 dB.

In Fig. 2 the frequency responses for the bands 250−2000 Hz, single reflection case, are shown. The differences in response across octaves can be attributed to the fact that if a ray is scattered, its direction will go randomly, causing more ray dispersion than from the vector based scattering in ODEON. Moreover, the scattering coefficient is frequency dependent.

[4] A simulation using an omni source was performed in ODEON and the resulting level difference was $\Delta L = -6.45$ dB which is closer to the result for the classical raytracing algorithm.

TABLE 3: ABSORPTION AND SCATTERING COEFFICIENTS USED IN THE SIMULATED SCENES. VALUES ARE PRESENTED AS *ABSORPTION* | *SCATTERING* IN %

Scene #	Material	63 Hz	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
2,3 & 5	*MDF 25 mm*	1.0 | 5.0	2.2 | 5.0	2.1 | 5.0	2.8 | 5.0	3.3 | 5.0	3.1 | 5.4	5.4 | 7.6	3.8 | 10.8
2,3 & 5	*100 % absorption*	1.0 | 0.0	1.0 | 0.0	1.0 | 0.0	1.0 | 0.0	1.0 | 0.0	1.0 | 0.0	1.0 | 0.0	1.0 | 0.0
5	*Tiles*	0.5 | 5.0	1.0 | 5.0	1.1 | 5.0	2.0 | 6.0	2.1 | 8.5	3.0 | 12.1	2.3 | 17.1	3.0 | 24.1
9	*Ceiling*	0.7 | 5.2	8.3 | 7.4	10.4 | 10.5	4.8 | 14.8	4.9 | 20.9	4.7 | 9.6	6.2 | 41.8	5.0 | 59.1
9	*Concrete*	7.7 | 5.0	8.5 | 5.0	7.5 | 5.0	5.6 | 5.0	5.9 | 6.0	5.9 | 8.5	4.4 | 12.1	4.1 | 17.1
9	*Floor*	8.6 | 5.0	7.1 | 5.0	9.1 | 5.0	7.0 | 5.0	6.5 | 5.0	6.2 | 6.6	4.3 | 9.4	3.3 | 13.2
9	*Plaster*	7.5 | 5.0	3.3 | 5.0	5.0 | 5.0	3.9 | 5.0	4.4 | 6.6	4.8 | 9.4	3.6 | 13.2	2.8 | 18.7
9	*Windows*	22.3 | 5.0	17.5 | 5.0	7.3 | 5.0	4.9 | 5.0	5.7 | 5.0	13.3 | 5.4	5.5 | 7.6	5.3 | 10.8

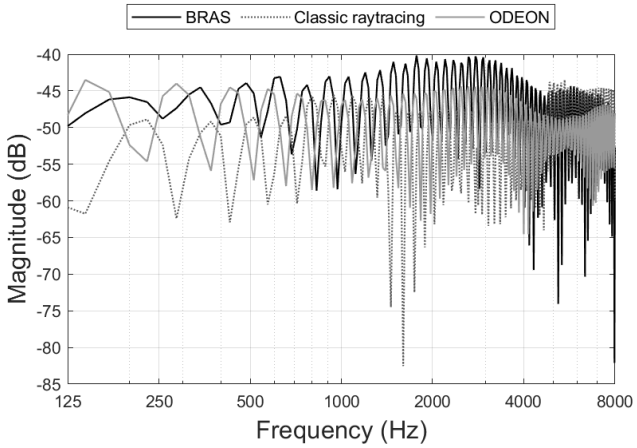

FIGURE 2: FREQUENCY RESPONSE OF THE IMPULSE RESPONSES FROM SCENE 2

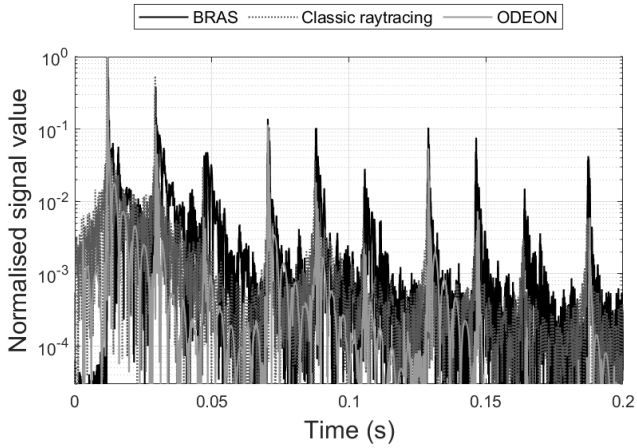

FIGURE 3: THE PROCESSED IMPULSE RESPONSES FROM SCENE 3

TABLE 4: TIME AND LEVEL DIFFERENCE FOR THE SIMULATED SINGLE REFLECTION SCENE COMPARED WITH THE REFERENCE

Algorithm	$FR - DS$ (ms) | Δt (ms) | %	ΔL (dB)
BRAS (reference)	6.848 | / | / |	/
Classic raytracing	6.875 | 0.027 | 0.39%	−4.66
ODEON	6.871| 0.023 | 0.33%	−1.35

Scene 3 - Multiple reflections

In Fig. 3 the simulated and measured IRs for *Scene 3*, multiple reflections case, are shown. In Fig. 4 the frequency responses of the compared signals are shown. As with the single reflection case the time arrival of the reflections are modelled accurately, but the level of the peaks and their decay tail does not align completely with the reference. Both the difference in source directivity and scatter modelling can be considered the reason for this outcome.

Similarly to the single reflection case, five simulations were performed with the classical raytracing method and the raw data prior to IR generation was inspected. The precise timing of raytraced contributions was extracted and their deviation from the average was computed. For these values, the Relative Standard Deviation (RSD), (σ / m.) was found to be fairly low, with a maximum at 0.18% for the first reflection and less than 0.1% for all other reflections. Within the time range of 200 ms the standard deviation for the five tests at each frequency band was below a tenth of a millisecond.

Regarding the robustness of the reflection levels, Fig. 5 shows the average difference in dB for each octave band of the reflections with respect to direct sound over time. It is shown that the standard deviation and standard error for the five tests slightly increase over time, from < 1 dB at 100 ms to < 4 dB at 300 ms. Fig. 6 shows in detail the variation of the standard deviation in dB over time, for the different frequency bands and the average[5]. Each reflection by the opposite plates causes geometrical dispersion and loss of rays, that are absorbed by the anechoic environment. For instance, of 2M rays shot, less than 2K are recorded by the receiver as direct sound, about a third of these captures the first reflection, about a sixth of this third the second and so on, with a progressive sparsity that soon causes more variability in the results.

Scene 5 - Diffraction

The results in Fig. 7 and 8 show that only the first order diffraction is modelled by ODEON. Due to the omission of the

[5]In the Figure two distinct events are plotted at 0.162 and 0.163 s, as well as 0.262 and 0.263 s. This is due to setting 1 ms time bins. Another consequence of this approach is that only the first 20 reflection events presented enough data at all frequency bands to be considered for this analysis.

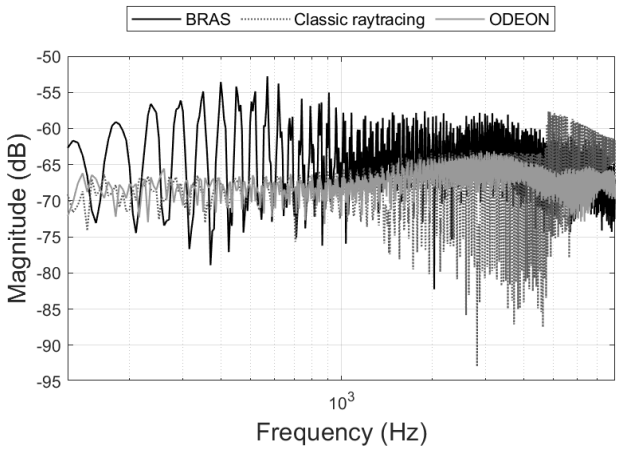

FIGURE 4: FREQUENCY RESPONSE OF CLASSICAL RAYTRACING AND ODEON COMPARED WITH THE MEASUREMENT IN SCENE 3

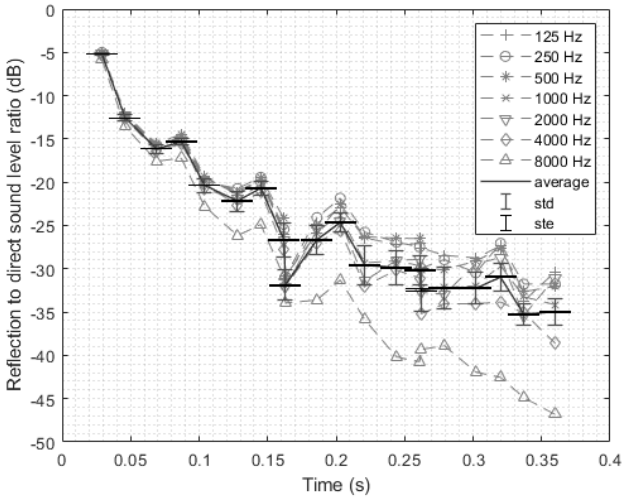

FIGURE 5: REFLECTION TO DIRECT SOUND RATIO IN SCENE 3 FOR FIVE SIMULATIONS

higher order diffraction contributions, the pattern in the frequency response is not captured. However, the frequency decay slope follows the reference.

Scene 9 - Small lecture room

Figure 9, 10 and 11 show the early and late IRs and frequency responses from *Scene 9* respectively. The solid lines in Fig.11 (gray for BRAS and black for the simulations) are the frequency responses smoothed to 1/3 octave bands.

In this scene, the arrival time of the direct sound for the simulated cases is less than the reference. For the classical raytracing the difference in air humidity and temperature contributes to this. When calculating the theoretical arrival time with respect to the source and receiver distance given in the dataset the arrival time is correct. However, the measurement was conducted with a loudspeaker array (low, mid, high) placed at different heights which then causes a slightly longer travelling distance at different frequencies.

In the early part of the IR ODEON captures the stronger reflections which are not present in the classical raytracing. Since the early part of the IR is partly modelled using image sources in ODEON, this is an outcome of the different algorithm designs. For the late part the classical raytracing has less energy in the decay tail whereas ODEON is closer to the measured reference.

In Tab. 5 the derived values for the reverberation time, T_{20}, are shown. The difference between the simulations and reference are shown by ΔT_{20} where a positive value indicates an overestimation and a negative value an underestimation of the reverberation time compared to the reference. The values are close to the measured reference with the exception of octave bands 250, 500 and 2000 Hz for the classical raytracing. For ODEON the deviation is less than 0.1 s in octave bands 250 and 4000 Hz. The global T_{20} which is the mean of the octave bands 250 – 4000 Hz matches well with the reference.

Five classical raytracing simulations were launched to assess the method's robustness. T_{20} values were extracted from both the processed IR, following the procedure described above, and from the raw raytracing results, following the procedure described in Ref. [40]. These results are presented in Table 6. It can be observed a difference above 0.1 s in the mean values[6] obtained from the two procedures only at 4 kHz. However, the standard deviation σ in seconds for the five T_{20} values calculated from the raw results is for all bands significantly lower than that from the processed IR T_{20}, difference visible also in the *RSD*. Hence, when using simulation data to train echolocation cues, it should be accounted that the IR generation process might introduce some additional uncertainties.

DISCUSSION

From the presented results it can be seen that even for simple controlled cases it is difficult to accurately represent a measured counterpart. There are several reasons for this, some of which have been briefly explained in the previous section.

The modelling of loudspeaker directivity has an impact on the level relationship of the reflections as seen in the simple scenes 2 and 3. In the case of multiple reflections more energy is transmitted to the panel located behind the loudspeaker and thereby an increase of level for the incoming reflections was expected in the simulations. An explanation can be that due to the random scattering in the classical raytracing, rays were lost in the anechoic environment and hence the level relation suffered. However, this should have happened with less evidence in ODEON, where vector based scattering and directive source is modelled.

[6]Mean T_{20} values (*m.*) where *IR* denote values that are derived from the processed impulse responses and *Raw* from the raw results.

TABLE 5: REVERBERATION TIME VALUES FOR SCENE 9 CALCULATED FROM IR

Impulse response origin	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	$T_{20,250-4000}$ (s)
BRAS	1.40	1.63	2.04	1.88	1.70	1.55	1.76
Classic raytracing	1.45	1.77	2.18	1.94	1.42	1.53	1.77
ODEON	1.53	1.65	2.15	2.03	1.82	1.65	1.86
Classic raytracing ΔT_{20}	0.07	0.15	0.14	0.06	-0.27	-0.02	0.01
ODEON ΔT_{20}	0.15	0.02	0.11	0.15	0.13	0.10	0.10

TABLE 6: REVERBERATION TIME ROBUSTNESS FOR CLASSICAL RAYTRACING (N=5)

	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	$T_{20,250-4000}$ (s)
m. Raw	1.45	1.60	2.15	1.94	1.38	1.70	1.75
m. IR	1.43	1.63	2.18	1.94	1.41	1.52	1.74
σ *Raw*	0.0085	0.0019	0.0156	0.0061	0.0021	0.0045	0.0032
σ *IR*	0.0744	0.0488	0.0911	0.0248	0.0252	0.0335	0.0206
RSD Raw	0.58%	0.12%	0.72%	0.31%	0.15%	0.27%	0.18%
RSD IR	5.21%	2.99%	4.18%	1.28%	1.79%	2.20%	1.19%

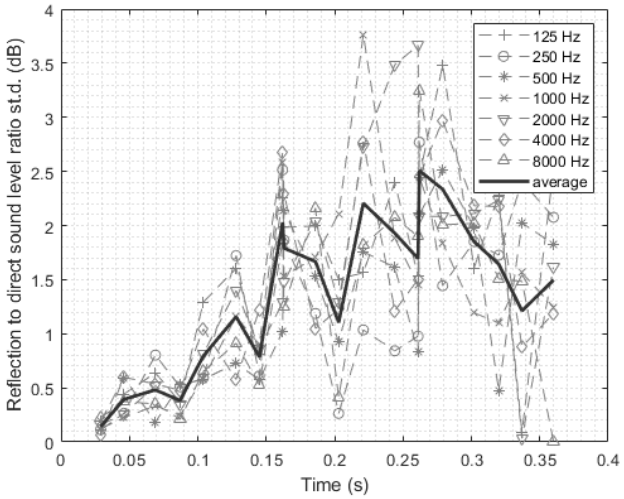

FIGURE 6: STANDARD DEVIATION OF REFLECTION TO DIRECT SOUND RATIO IN SCENE 3 FOR FIVE SIMULATIONS

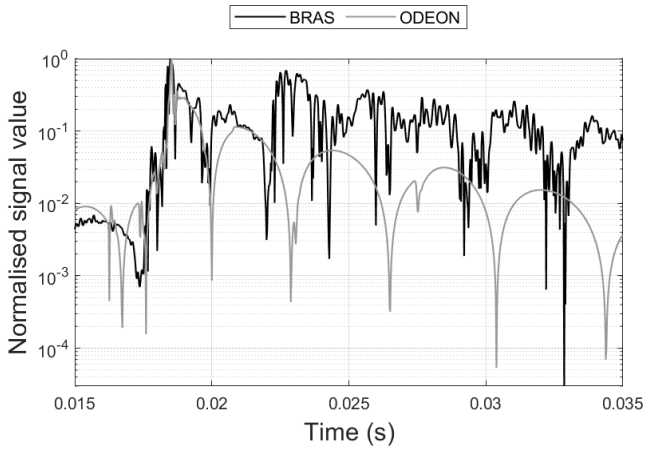

FIGURE 7: IMPULSE RESPONSES IN SCENE 5

Another reason for the difference in level relation is due to the absorption and scattering coefficients which in reality are angle dependant. For this study, the coefficients were provided (normal incidence for *Scene 2*, *3* and *5* and random incidence deriving from Eyring fitting for *Scene 9*) and used up to 8000 Hz. Furthermore, a limitation in practice is that absorption values are often provided in the frequency range of 125 – 4000 Hz and hence for calculation of frequencies beyond that limit the calculation software need to extrapolate or take this into account in some other way. However, for deriving room acoustical parameters in room acoustic design these are proven to be sufficient which is also seen for the classic raytracing approach where T_{20} aligned with the reference values.

Phase information is part of the image source method and ODEON utilises this in the early part of the IR. However, in the late part as well as for classic raytracing in general, the phase is random and thereby not accurately modelled.

The choice of adopting the *Surround* option from ODEON follows the need of retrieving a monaural IR that could be used for comparison. It was observed that beside an artificial delay, the *Surround* wave[7] file contained also a different frequency response when compared to W channel. It was later found that a specific set of HRTFs provided with the software package could have been used to simulate monoaural receivers when using binaural RIRs. A follow-up study would help clarifying the impact of the auralisation process in the IR generation prior to moving to the study of binaural RIRs.

Therefore, future studies that adopt GA algorithms to generate mono or binaural RIRs should take into account the variability of the simulated outcomes with respect to frequency content and relative level of reflections. In the study of human echolocation

[7]File format .wav.

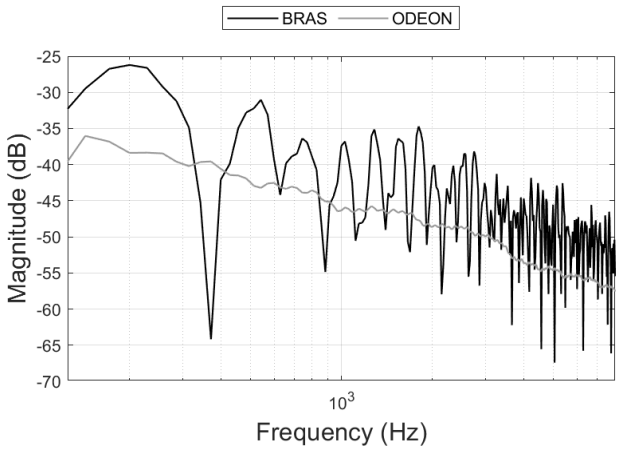

FIGURE 8: FREQUENCY RESPONSE FROM THE IMPULSE RESPONSES IN SCENE 5

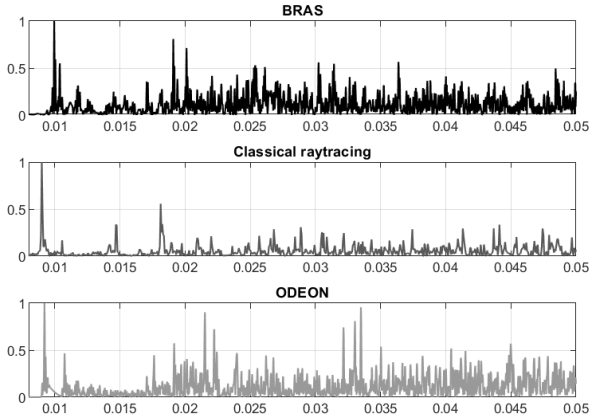

FIGURE 9: EARLY PART OF THE IMPULSE RESPONSES IN SCENE 9

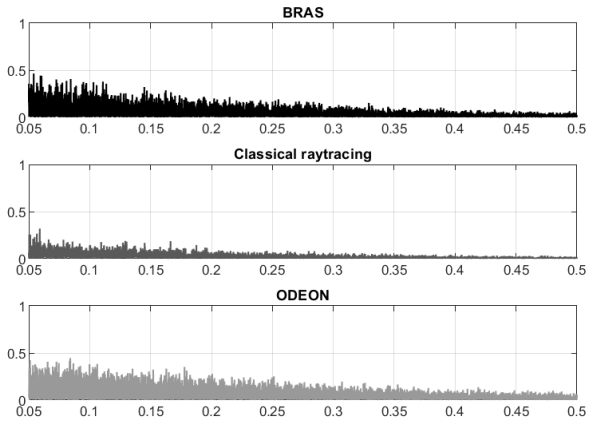

FIGURE 10: LATE PART OF THE IMPULSE RESPONSES IN SCENE 9

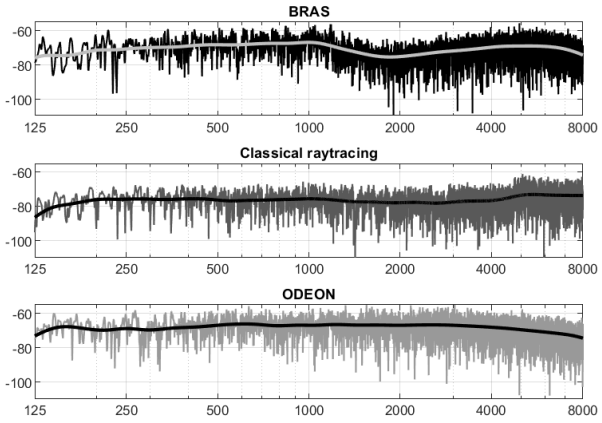

FIGURE 11: FREQUENCY RESPONSE FROM THE IMPULSE RESPONSES IN SCENE 9

cues and spatial reconstruction capabilities, this appears to be a crucial condition to be met prior to looking at binaural RIR standard descriptors.

CONCLUSION

In the tested GA calculation approaches, the arrival time of the reflections are simulated with precision for the simple controlled cases, but for complex rooms with many reflecting surfaces and different surface properties the tested methods diverge from the measured reference. Furthermore, due to the many uncertainties implied in adopting angle dependent or fitted absorption and scattering coefficients and the implementation of scattering and diffraction in the simulation, the level relations between the reflections seem to be too easily disrupted within the expected sensitivity of trained echolocators. With the aim of a very precise simulated representation of a room, GA approaches suffer from the limitations set by the actual algorithms as well as the input data fed into the simulation.

ACKNOWLEDGMENTS

This research is funded through the Science Park of University of Iceland fund and Treble Technologies.

REFERENCES

[1] "Visioneers – a division of World Access for the Blind." URL https://www.visioneers.org/. Accessed on 2022-08-03.

[2] Ekkel, M. R., Van Lier, R. and Steenbergen, B. "Learning to echolocate in sighted people: a correlational study on attention, working memory and spatial abilities." *Experimental Brain Research* Vol. 235 No. 3 (2017): pp. 809–818. DOI 10.1007/s00221-016-4833-z. URL https://dx.doi.org/10.1007/s00221-016-4833-z.

[3] Holmes, Nicole. "An Echolocation Training Package." *International Journal of Orientation & Mobility* Vol. 4 No. 1.

[4] Brunes, Audun, B. Hansen, Marianne and Heir, Trond. "Loneliness among adults with visual impairment: preva-

lence, associated factors, and relationship to life satisfaction." *Health and Quality of Life Outcomes* Vol. 17 No. 1. DOI 10.1186/s12955-019-1096-y. URL https://dx.doi.org/10.1186/s12955-019-1096-y.

[5] Brunes, Audun and Heir, Trond. "Serious Life Events in People with Visual Impairment Versus the General Population." *International Journal of Environmental Research and Public Health* Vol. 18 No. 21 (2021): p. 11536. DOI 10.3390/ijerph182111536. URL https://dx.doi.org/10.3390/ijerph182111536.

[6] Thaler, L. and Goodale, M. A. "Echolocation in humans: an overview." *Wiley Interdiscip Rev Cogn Sci* Vol. 7 No. 6 (2016): pp. 382–393. DOI 10.1002/wcs.1408. URL https://www.ncbi.nlm.nih.gov/pubmed/27538733.

[7] Worchel, Philip, Mauney, Jack and Andrew, John G. "The perception of obstacles by the blind." *Journal of Experimental Psychology* Vol. 40 No. 6 (1950): pp. 746–751. DOI 10.1037/h0060950.

[8] Kolarik, A. J., Cirstea, S., Pardhan, S. and Moore, B. C. "A summary of research investigating echolocation abilities of blind and sighted humans." *Hear Res* Vol. 310 (2014): pp. 60–8. DOI 10.1016/j.heares.2014.01.010. URL https://www.ncbi.nlm.nih.gov/pubmed/24524865.

[9] Schenkman, Bo N. "Identification of ground materials with the aid of tapping sounds and vibrations of long canes for the blind." *Ergonomics* Vol. 29 No. 8 (1986): pp. 985–998. DOI 10.1080/00140138608967212. URL https://dx.doi.org/10.1080/00140138608967212.

[10] Au, Whitlow W. L. and Martin, Douglas W. "Insights into dolphin sonar discrimination capabilities from human listening experiments." *The Journal of the Acoustical Society of America* Vol. 86 No. 5 (1989): pp. 1662–1670. DOI 10.1121/1.398596. URL https://dx.doi.org/10.1121/1.398596.

[11] Ashmead, D. H. and Wall, R. S. "Auditory perception of walls via spectral variations in the ambient sound field." *Journal of Rehabilitation Research and Development* Vol. 36 No. 4 (1999): pp. 313–322.

[12] Doucet, M. E., Guillemot, J. P., Lassonde, M., Gagné, J. P., Leclerc, C. and Lepore, F. "Blind subjects process auditory spectral cues more efficiently than sighted individuals." *Experimental Brain Research* Vol. 160 No. 2 (2005): pp. 194–202. DOI 10.1007/s00221-004-2000-4. URL https://dx.doi.org/10.1007/s00221-004-2000-4.

[13] Voss, Patrice, Lassonde, Maryse, Gougoux, Frederic, Fortin, Madeleine, Guillemot, Jean-Paul and Lepore, Franco. "Early- and Late-Onset Blind Individuals Show Supra-Normal Auditory Abilities in Far-Space." *Current Biology* Vol. 14 No. 19 (2004): pp. 1734–1738. DOI 10.1016/j.cub.2004.09.051. URL https://dx.doi.org/10.1016/j.cub.2004.09.051.

[14] Muchnik, Chava, Efrati, Michal, Nemeth, Esther, Malin, Michal and Hildesheimer, Minka. "Central Auditory Skills in Blind and Sighted Subjects." *Scandinavian Audiology* Vol. 20 (2009): pp. 19–23. DOI 10.3109/01050399109070785.

[15] Thaler, L. and Castillo-Serrano, J. "People's Ability to Detect Objects Using Click-Based Echolocation: A Direct Comparison between Mouth-Clicks and Clicks Made by a Loudspeaker." *PLoS One* Vol. 11 No. 5 (2016): p. e0154868. DOI 10.1371/journal.pone.0154868. URL https://www.ncbi.nlm.nih.gov/pubmed/27135407.

[16] Teng, S., Puri, A. and Whitney, D. "Ultrafine spatial acuity of blind expert human echolocators." *Exp Brain Res* Vol. 216 No. 4 (2012): pp. 483–8. DOI 10.1007/s00221-011-2951-1. URL https://www.ncbi.nlm.nih.gov/pubmed/22101568.

[17] Cotzin, Milton and Dallenbach, Karl M. ""Facial Vision:" The Role of Pitch and Loudness in the Perception of Obstacles by the Blind." *The American Journal of Psychology* Vol. 63 No. 4 (1950): p. 485. DOI 10.2307/1418868. URL https://dx.doi.org/10.2307/1418868.

[18] Rowan, Daniel, Papadopoulos, Timos, Edwards, David, Holmes, Hannah, Hollingdale, Anna, Evans, Leah and Allen, Robert. "Identification of the lateral position of a virtual object based on echoes by humans." *Hearing Research* Vol. 300 (2013): pp. 56–65. DOI 10.1016/j.heares.2013.03.005. URL https://dx.doi.org/10.1016/j.heares.2013.03.005.

[19] DeLong, C. M., Au, W. W. and Stamper, S. A. "Echo features used by human listeners to discriminate among objects that vary in material or wall thickness: implications for echolocating dolphins." *J Acoust Soc Am* Vol. 121 No. 1 (2007): pp. 605–17. DOI 10.1121/1.2400848. URL https://www.ncbi.nlm.nih.gov/pubmed/17297814.

[20] Hausfeld, Steven, Power, P. Roderick, Gorta, Angela and Harris, Patricia. "ECHO PERCEPTION OF SHAPE AND TEXTURE BY SIGHTED SUB JECTS." *Perceptual and Motor Skills* No. 55 (1982): pp. 623–632.

[21] Seki, Y. and Sato, T. "A training system of orientation and mobility for blind people using acoustic virtual reality." *IEEE Trans Neural Syst Rehabil Eng* Vol. 19 No. 1 (2011): pp. 95–104. DOI 10.1109/TNSRE.2010.2064791. URL https://www.ncbi.nlm.nih.gov/pubmed/20805059.

[22] Picinali, Lorenzo, Afonso, Amandine, Denis, Michel and Katz, Brian FG. "Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge." *International Journal of Human-Computer Studies* Vol. 72 No. 4 (2014): pp. 393–407.

[23] Schroeder, MR, Atal, BS and Bird, C. "Digital computers in room acoustics." *Proc. 4th ICA, Copenhagen M* Vol. 21.

[24] Sorsdal, A. Krokstad; S. Strom; S. "CALCULATING THE ACOUSTICAL ROOM RESPONSE BY THE USE OF A RAY TRACING TECHNIQUE." *Journal of Sound and Vibration* Vol. 8 No. 1 (1968): pp. 118–125.

[25] Vorländer, Michael. *Auralization - Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, second edition ed. ASA Press (2020).

[26] Allen, Jont B and Berkley, David A. "Image method for efficiently simulating small-room acoustics." *The Journal of the Acoustical Society of America* Vol. 65 No. 4 (1979): pp. 943–950.

[27] Naylor, Graham. "Treatment of early and late reflections in a hybrid computer model for room acoustics." *The Journal of the Acoustical Society of America* Vol. 92 No. 4 (1992): pp. 2345–2345. DOI 10.1121/1.404930.

[28] Rindel, Jens Holger. "The Use of Computer Modeling in Room Acoustics." *Journal of Vibroengineering* No. 3 (2000): pp. 219–224.

[29] A/S, Odeon. "ODEON User's manual." (2021). URL https://odeon.dk/download/Version17/OdeonManual.pdf. Accessed on 2022-08-03.

[30] Zeng, Xiangyang, Christensen, Claus Lynge and Rindel, Jens Holger. "Practical methods to define scattering coefficients in a room acoustics computer model." *Applied Acoustics* Vol. 67 No. 8 (2006): pp. 771–786. DOI 10.1016/j.apacoust.2005.12.001.

[31] Jeong, C. H. and Ih, J. G. "Effects of source and receiver locations in predicting room transfer functions by a phased beam tracing method." *J Acoust Soc Am* Vol. 131 No. 5 (2012): pp. 3864–75. DOI 10.1121/1.3699268. URL https://www.ncbi.nlm.nih.gov/pubmed/22559362.

[32] Savioja, L. and Svensson, U. P. "Overview of geometrical room acoustic modeling techniques." *J Acoust Soc Am* Vol. 138 No. 2 (2015): pp. 708–30. DOI 10.1121/1.4926438. URL https://www.ncbi.nlm.nih.gov/pubmed/26328688.

[33] Aspöck, Lukas, Brinkmann, Fabian, Ackermann, David, Weinzierl, Stefan and Vorländer, Michael. "BRAS - Benchmark for Room Acoustical Simulation." DOI 10.14279/depositonce-6726.3. URL http://dx.doi.org/10.14279/depositonce-6726.3.

[34] Brinkmann, Fabian, Aspöck, Lukas, Ackermann, David, Opdam, Rob, Vorländer, Michael and Weinzierl, Stefan. "A benchmark for room acoustical simulation. Concept and database." *Applied Acoustics* Vol. 176. DOI 10.1016/j.apacoust.2020.107867.

[35] Kulowski, Andrzej. "Algorithmic representation of the ray tracing technique." *Applied Acoustics* Vol. 18 No. 6 (1985): pp. 449–469.

[36] Aspöck, Lukas, Brinkmann, Fabian, Ackermann, David, Weinzierl, Stefan and Vorländer, Michael. "BRAS - Benchmark for Room Acoustical Simulation." (2020). DOI 10.14279/depositonce-6726.3. URL http://dx.doi.org/10.14279/depositonce-6726.3.

[37] Harris, Cyril M. "Absorption of sound in air versus humidity and temperature." *The Journal of the Acoustical Society of America* Vol. 40 No. 1 (1966): pp. 148–159.

[38] Brinkmann, F., Aspock, L., Ackermann, D., Lepa, S., Vorlander, M. and Weinzierl, S. "A round robin on room acoustical simulation and auralization." *J Acoust Soc Am* Vol. 145 No. 4 (2019): p. 2746. DOI 10.1121/1.5096178. URL https://www.ncbi.nlm.nih.gov/pubmed/31046379.

[39] Berzborn, Marco, Bomhardt, Ramona, Klein, Johannes, Richter, Jan-Gerrit and Vorländer, Michael. "The ITA-Toolbox: An open source MATLAB toolbox for acoustic measurements and signal processing." *43th Annual German Congress on Acoustics, Kiel (Germany)*, Vol. 6: pp. 222–225. 2017.

[40] ISO 3382-1. "Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces." Standard ISO/TC 43/SC 2 3382-1:2009. International Organization for Standardization, Switzerland. 2009. URL https://www.iso.org/standard/40979.html.

[41] Alpkocak, Malik, Adil; Kemal. "Computing Impulse Response of Room Acoustics Using the Ray-Tracing Method in Time Domain." *Archives of Acoustics* DOI 10.2478/v10168-010-0039-8.

[42] Kuttruff, Heinrich. *Room Acoustics*, fifth edition ed. Spon press (2009).

[43] Papadopoulos, Timos, Edwards, David S., Rowan, Daniel and Allen, Robert. "Identification of auditory cues utilized in human echolocation—Objective measurement results." *Biomedical Signal Processing and Control* Vol. 6 No. 3 (2011): pp. 280–290. DOI 10.1016/j.bspc.2011.03.005.

[44] Schenkman, Bo N. and Gidla, Vijay Kiran. "Detection, thresholds of human echolocation in static situations for distance, pitch, loudness and sharpness." *Applied Acoustics* Vol. 163. DOI 10.1016/j.apacoust.2020.107214.

[45] Wallmeier, L. and Wiegrebe, L. "Self-motion facilitates echo-acoustic orientation in humans." *R Soc Open Sci* Vol. 1 No. 3 (2014): p. 140185. DOI 10.1098/rsos.140185. URL https://www.ncbi.nlm.nih.gov/pubmed/26064556.