Università Politecnica delle Marche
Scuola di Dottorato di Ricerca in Scienze dell'Ingegneria
Corso di Dottorato in Ingegneria Industriale

Auralization of noise sources to support interior car sound design

Ph.D. Dissertation of:

Mateusz Matuszewski

Supervisor:

Prof. Paolo Castellini

Assistant Supervisor:

Dr Paolo Chiariotti

Dr Karl Janssens

Ph.D. Course coordinator:

Prof. F. Mandorli

XV edition - new series

Università Politecnica delle Marche
Dipartimento di (nome del dipartimento dove la tesi e' stata sviluppata)
Via Brecce Bianche — 60131 - Ancona, Italy

Acknowledgements

This thesis would not have possible without the help and contributions of a number of people. I would like to express my gratitude to all people involved in the creation of the EID project ENHANCED in which I had a pleasure to participate: my supervisor Prof. Paolo Castellini, university assistant supervisors Prof. Milena Martarelli and Dr Paolo Chiariotti and my industry assistant supervisor Dr Karl Janssens. I sincerely believe that all of you had the greatest impact on my career and personal development. The ENHANCED project gave a unique opportunity to learn new technical skills, broaden my horizons, meet great people and last but not least, to write this thesis. Thank you very much for this fruitful time spent together.

My additional sincere thanks also goes to Dr Paolo Chiariotti for his constant encouragement and energetic motivation. On a daily basis, I was fortunate to count on his help in every matter. He gave me a wonderful support by teaching me new technical skills and by reviewing and criticising my work. His priceless help was crucial for this thesis.

I would also thank to reviewers Prof. Toon van Waterschoot from KU Leuven and Prof. Etienne Parizet from INSA Lyon for their insightful comments on this thesis.

At the end, I would like to thank my colleagues from UNIVPM and Siemens PLM, especially to Claudio Colangeli and Giampiero Accardo for all the good times shared in these last three years.

Abstract

Current trend in the automotive industry is to replace “traditional” internal combustion engines into electrical units. As a result, cars are becoming quieter, which brings sound sources originally masked by combustion engines to the top. One of the most industrially important “new” sources is wind noise. Performing investigation on sounds, especially on those which cannot be measured separately, require sophisticate methods, like auralization. This thesis is mainly focused on the development of a new time-saving auralization method, tailored for car interior applications, which preserve sufficient fidelity in terms of wind noise reproduction. The proposed method aims to make it possible to auralize wind noise without any prior knowledge of the geometry of the car under investigation. This goal is achieved by using a statistical model of key Room Impulse Response (RIR) parameters and early reflections. Quality of sound auralization highly depends on Head-Related Transfer Functions (HRTFs) used. A new approach for measuring and processing HRTFs targeted at auralizing sound in car cabins is proposed. The remaining question is to understand how many reflections are effectively needed for providing a realistic listening experience. Therefore Investigations on Room Impulse Response modifications are included in the thesis. Also a human-oriented approach to define the minimum accuracy required for source localization techniques when used in car interiors is presented. By exploiting sound source localization and separation it becomes possible to virtually relocate sources and to perform what-if analyses. The proposed approach for in-vehicle auralization can be a substantial economic benefit for car manufactures considering the high costs of wind tunnel tests.

Contents

List of Figures

List of Tables

List of Acronyms

BRIR	Binaural Room Impulse Response
DRR	Direct to Reverberation Ratio
EDC	Energy Decay Curve
ELR	Early to Late reverberation Ratio
ETD	Early Decay Time
FEM	Finite Element Method
FFT	Fast Fourier Transform
GA	Geometrical Acoustics
HATS	Head and Torso Simulator
HRTF	Head Related Transfer Function
IFFT	Inverse Fast Fourier Transform
ILD	Interaural Level Difference
ITD	Interaural Time Difference
RIR	Room Impulse Response
RT	Ray Tracing
RT	Reverberation Time
RTF	Room Transfer Function
SNR	Signal to Noise Ratio
SSL	Sound Source Localization

General introduction and thesis outline

Cars, while in use, will always generate noise. In most of the cases noise is an undesired consequence of driving a car. However, nothing related to humans is straightforward. Perception of acoustics events are extremely subjective. For someone, a loud, sportive car can be, acoustically speaking, an ideal one, while for others it can be extremely annoying. Despite of this trivial example, it is well recognized by everyone that most of the car noises are unpleasant and this explains the current challenges that NVH engineers are asked to face: to optimize the passengers' acoustic experience in car cabins. Traditionally, there are three main types of noise sources in car cabin interior acoustics.. The first group relates to tire noise, because it is caused by the interaction between tire and pavement. The second, the most characteristic in car interior acoustic, is the powertrain noise. This noise is usually dominated by harmonics that are related to the engine orders and/or to gearbox rattling. The last group addresses wind noise; this noise is caused by air turbulences over the car body. All the three classes are highly related to car speed. Indeed, at low speed, Engine noise is the most dominant noise source. While the car speed increases, tire noise and wind noise becomes more dominant. At highway speeds, wind noise is critical in order to meet high quality customers expectations.

Current trend in the automotive industry is to replace "traditional" internal combustion engines into electrical units. As a result, cars are becoming quieter, which brings sound sources originally masked by combustion engines to the top. One of the most industrially important "new" sources is wind noise. Therefore, in the thesis, this type of source will be used as an input for further analysis.

Performing investigation on sounds, especially on those, which cannot be measured separately, require sophisticate methods, like auralization. With this approach, it is possible to predict the character of sound signals that are generated at the source and modified by reinforcement, propagation and transmission in systems such as car interiors [1].

This thesis is mainly focused on the development of a new time-saving auralization method, tailored for car interior applications, which preserve sufficient fidelity in terms of sound reproduction. The proposed method aims at making it possible to auralize wind noise without any prior knowledge of the geometry of the car under investigation. This goal is achieved by using a statistical model of key Room Impulse Response (RIR) parameters and early reflections. Additionally, by exploiting sound source localization and separation techniques (e.g. time domain beamforming and Transfer Path Analysis) it becomes possible to virtually relocate sources and to perform what-if analyses. Currently, in order to "turn off" wind noise sources like side mirrors and evaluate the influence of this part to the overall noise, it is

necessary to dismount it from the vehicle and to perform measurement in a wind tunnel again. The proposed approach can be a substantial economic benefit for car manufactures considering the high costs of wind tunnel tests.
The block diagram in Fig.0-1 presents an outline of the overall scope of the thesis. Each block within the dashed area indicates a research topic carried on during the PhD activity. These topics will be presented and discussed more in detail in dedicated chapters along this thesis. The three blocks on top represent possible links with other approaches that can produce inputs for the techniques presented within this thesis.

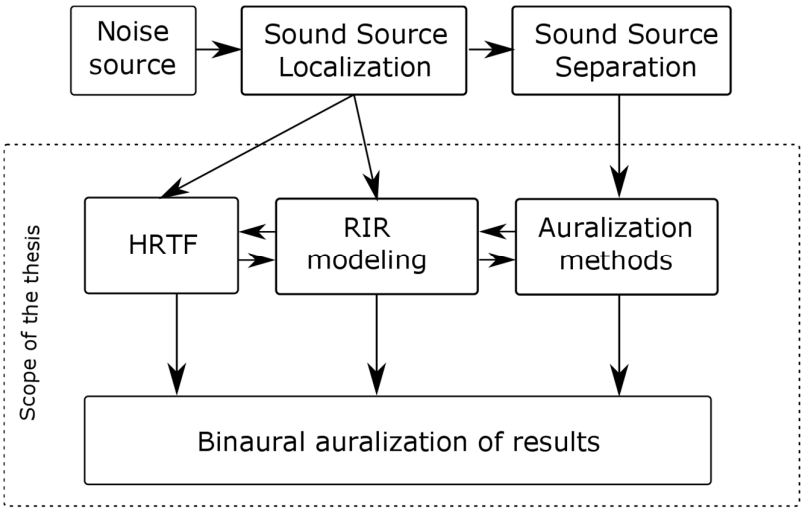


Fig.0-1 Block diagram presenting the outline of the thesis. Blocks inside the dashed block indicate investigated areas.

Chapter 1 gives the theoretical foundations of car interior acoustics, focusing mostly on the structure and on parametrization methods of Impulse Responses. Relevant auralization techniques are also presented. This chapter gives an overview of sources contributing to wind noise in car cabins.

Chapter 2 presents investigations of driver's position influence on Head Related Transfer Function

Chapter 3 New auralization methods require investigations on human perception when related to acoustics input estimation techniques (e.g. beamforming measurements). This Chapter aims at defining the optimal number of reflections in RIRs that is necessary to use for in-vehicle sound auralization purposes.

Chapter 4 aims at presenting a human-oriented approach to define the minimum accuracy required for source localization techniques when used in car interiors in combination with auralization strategies.

Chapter 5 describes an auralization approach based on a statistical model. The main hypothesis on which the method grounds is that, in car cabins, the diversity of materials and of possible driver's position, causes the approximation of the acoustic environment for frequencies far above the Schroeder frequency. This holding, the method proposed relays on a statistical approach linking RIR parameters, direct path, early and late reflections. A detailed description of the model implementation and input parameters estimation is provided in Chapter 5.

Chapter 1

Car cabin acoustics

1.1 Room Impulse Response

In car cabins, like in all enclosed spaced, whenever a sound source radiates, a signal recorded by the receiver not only consists of the excitation by the source. In fact, the propagating sound waves are reflected on the car panels, diffracted on edges, scattered on surface textures, absorbed by surface materials, and attenuated by air, until they finally reach the receiver. Thus, the total sound perceived by a listener consists of a direct sound and a series of delayed and attenuated reflections that arrive shortly after the direct sound (commonly called *early reflections*) and reflections that arrive after the early reflections (commonly called *late reverberation*). The latter events describe the acoustical behaviour of the cavity. This process is illustrated in Fig. 1-1. When the source signal is a single pulse, acoustic response of the cavity can be represented by so-called Room Impulse Response (RIR) and denoted by *h(t)*. The Fourier transform of the RIR is typically called the Room Transfer Function (RTF) and denoted by *H(f)*.

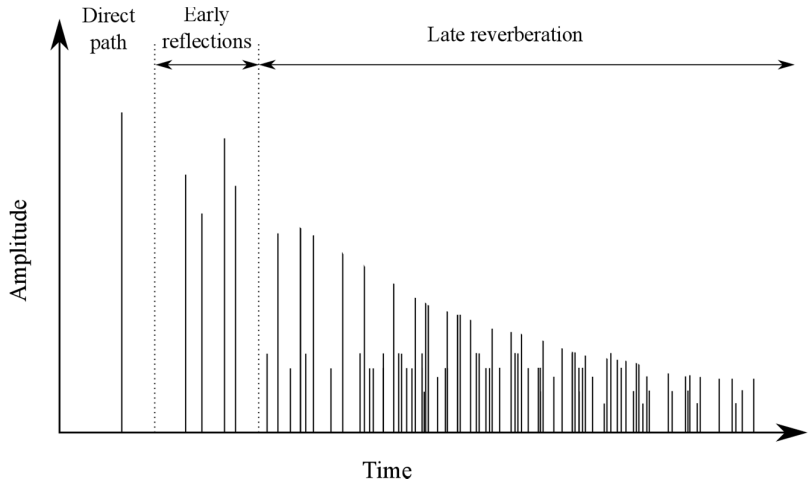


Fig. 1-1 An illustration of a RIR division into three parts: direct path, early reflections and late reverberation

Below, a short discussion of the three main RIR components will be presented.

Direct sound. It is the first sound that reaches the receiver without any reflection. Considering a wind noise, the main noise source considered in this thesis, the source and the receiver are always in line of sight. It means that in all cases the direct sound is present. This information is of great importance, because the direction of arrival of the direct sound corresponds to the perception of the entire sound localized. This is often referred to as the *precedence effect* [2]. In case of non-line of sight conditions, there is no direct sound present. The time delay between the initial excitation of the source and the actual time of direct sound arrival is dependent on the distance and on the velocity of the sound. In car cabins the time delay usually does not exceed 4ms.

A short time after the direct sound first, sounds which were reflected from doors, windshield and dashboard are received. These are called **early reflections** and they are separated in both time and direction from the direct sound. Amplitudes and direction of arrivals of early reflections will vary as the source or the receiver moves inside the car cabin. Early reflections play significant role in perception of a sound in car cabins, since reflecting surfaces are very close to the receiver. This attribute can cause some perceptual phenomena like *image shift* or *tone coloration* [3, 4]. Chapter 3 will present more detailed investigations on how early reflections are perceived by a listener.

Late reverberation consist of reflections that arrive with larger delays than the direct sound and early reflections. After a certain period of time, which is, in car cabins, typically 10-15ms after the direct sound, the number of reflections increases gradually and the human hearing cannot anymore perceive them as single events. Late reverberation are nearly independent of the listener's position, as the human hearing starts to perform a quite rough energetic integration over a certain time slot and angle field [5].

1.2 Methods for IR calculation

1.2.1 Ray tracing method

Ray Tracing (RT) is a very powerful Geometrical Acoustics (GA) method. In RT, a sound source at a given position emits numerous sound particles in all directions and each sound particle is tracked around the room. The principle of a particle propagation is considered in much the same way as light rays are treated in optics. Each particle travels on a straight path until it hits a wall loosing energy. After every reflection, a new direction of propagation is determined according to Snell's law. When a particle reaches a receiver position, the energy and time of arrival are stored and can be presented in an echogram. It is important to define an area or a volume around the receiver in order to catch particles when passing by. RT,

generally speaking, is a sort of Monte Carlo simulation. Fig. 1-2 illustrates appended rays propagation paths obtained by the RT method.

First works on ray acoustics were already known in the early 1930s [6]. Many years later, with the arrival of appropriate digital computing systems, in [7] was presented the first computer based RT technique, whose aim was the study of the envelope of early reflected sound energy in time and space taking into account the specific shape of a room. In 1970 Schroeder presented a complete auralization framework for the generation and playback of computer aided room acoustic simulations based on the ray tracing technique [8]. From then on, many works related to this technique have been presented to the scientific community [9, 10, 11, 12]. RT is also implemented in several commercially available room acoustic simulation software.

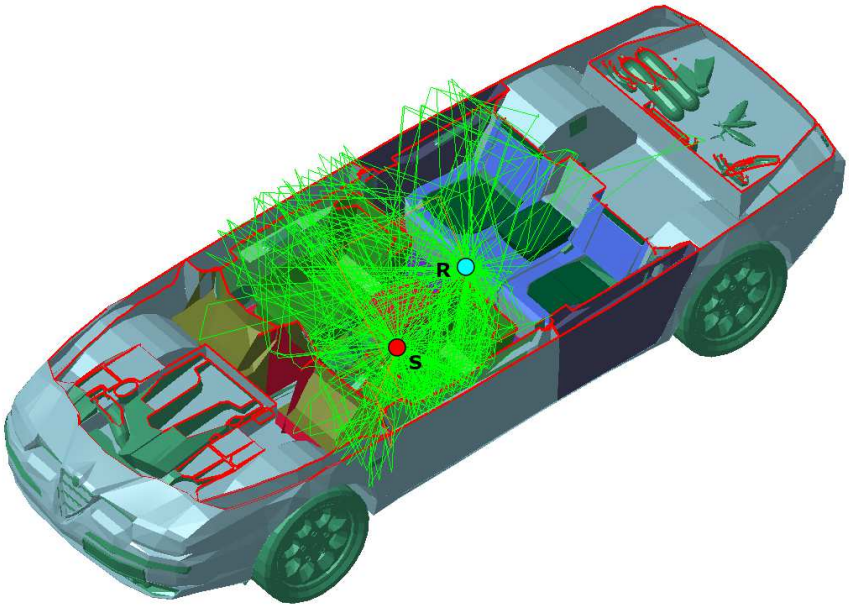


Fig. 1-2 Cross section of a car illustrating source propagation inside a car cabin calculated using RT method. Green lines indicate sound paths from a source (S) to a receiver (R).

1.2.2 Statistical methods

Ray tracing method and wave based methods of calculating RIRs have common big disadvantage. They all require a car model and a set of acoustic parameters that characterize

the boundary. Similar results might be achieved in much simpler way. Statistical Room Acoustics (SRA), under certain assumptions, can provide a statistical description of RTF between source and receiver on the basis of few parameters like source-receiver distance, source location, cavity volume etc.

Sabine presented the earliest attempt to room statistical methods. He introduced, in the late 19th, a calculation method of the reverberation time of a space without considering the details of its geometry. More than 50 years later, Schroeder extended Sabine's fundamental work [13] and derived a set of statistical properties describing the frequency response of a random impulse response. In [14] Moorer et al. noted that the impulse responses in the finest concert halls around the world sounded remarkably similar to white noise with an exponential amplitude envelope. To test this observation, they generated synthetic impulse responses by shaping unit-variance Gaussian pseudo-random sequences with an exponential of the desired length. The direct sound was added by including an impulse at the beginning. Later, Polack [15] developed a time-domain model excluding the contribution of the direct path and describing RIR as a realization of a non-stationary stochastic process:

$$h(t) = \begin{cases} b(t)e^{-\zeta t}, & t \geq 0; \\ 0, & t < 0. \end{cases} \qquad 1.1$$

where *b(t)* is a zero-mean stationary Gaussian noise, and ζ is a damping constant related to the reverberation time, RT60, by

$$\zeta = \frac{\log(10^3)}{T_{60}} \quad . \qquad 1.2$$

This statistical model is valid when the sound field is diffuse, i.e., when the sound energy density and the direction of the intensity vector are uniformly distributed across the room. The peaks in the acoustic impulse response then no longer correspond to the arrivals of individual reflections. The point at which RIR can be adequately approximated using statistical methods is called the *mixing time*. For rooms, fairly arbitrary value of 80 ms is usually taken. In car acoustics, the mixing time is much shorter and varies from 10 ms to 20 ms. The simplest method for calculating the mixing time is based on the average number of reflections, i.e., the echo density (D_e), which is approximately proportional to the square of time *t* [16]:

$$D_{\mathrm{e}}(t) \approx 4\pi c^3 \frac{t^2}{V}, \qquad 1.3$$

where V is room volume in m^3 and c denotes speed of sound in m/s. Modifying the equation above and taking four reflections as suggested in [17], the popular estimation of the mixing time equals:

$$t_{mix} \approx \sqrt{V} \quad \text{in ms.} \qquad 1.4$$

For car interior acoustics, this method provides a mixing time of about 2 ms, which is rather incorrect. Literature gives couple of different methods for the mixing time estimation [17, 18, 19, 20, 21, 22]. All of them are compared in [23]. Among them, the method proposed in [22] based on kurtosis calculations seems to be the most suitable for car RIR. Later in Chapter 3, it will be shown that the mixing time can be perceptually estimated based on the detection of a threshold on early reflections in RIR.

Recently, two modifications of the Polack's model have been introduced. A generalized Polack model, which is valid when the source-receiver distance is smaller than the critical distance, was presented in [24]. In [25] a model including early decay time as a second decay curve was also developed.

1.3 RIR parametrization

1.3.1 Reverberation Time

Reverberation time (RT_{60}) is the well-known, and probably the most common, parameter describing the acoustic behaviour of an acoustic cavity. This term was introduced in Sabine's pioneering research, in which he noticed that reverberation time was proportional to the volume of the room and inversely proportional to the amount of absorption. Because the absorptive properties of materials vary as a function of frequency, the reverberation time is also frequency dependent.

Reverberation time, is calculated from the energy-decay curve (EDC) of the RIR, which can be obtained by integrating the impulse response h of the cavity as follows [26]

$$EDC(t) = \int_t^\infty h^2(\tau)\, d\tau \quad . \qquad 1.5$$

The EDC(t) is the amount of energy remaining in the IR, h, at time t. If the Signal to Noise Ratio (SNR) is sufficient, RT_{60} usually refers to the measured time for the EDC to decay

60dB. In practice very often RT_{60} is estimated based on the lower decay, e.g. 30dB as illustrated in Fig. 1-3

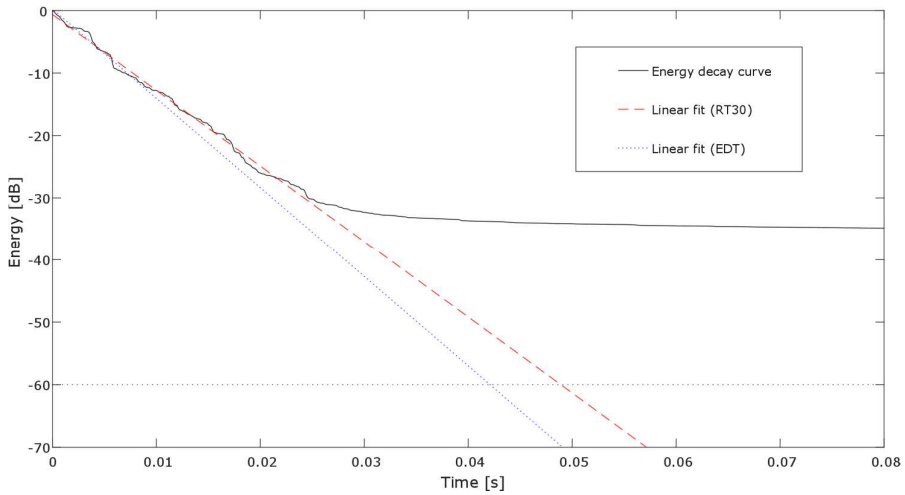


Fig. 1-3 Illustration of how reverberation time and early decay time are calculated from an energy decay curve.

1.3.2 Early Decay Time

Early reflections may decay with different rate than the late part of RIR. Jordan [27] proposed a parameter, very similar to RT, called Early Decay Time (EDT). The difference relies on the decay of EDC at which EDT is estimated: 10 dB instead of 60dB.
Fig. 1-3 presents an estimation method to derive the EDT (blue dotted line) and the RT (red line).

1.3.3 Direct-to-Reverberant Energy Ratio

Direct to Reverberant Radio (DRR) is an objective measure of RIR describing the energy of a direct sound with respect to all reflections. If the *h(t)* is a RIR, the Direct-to-Reverberant Ratio (DRR) is denoted as

$$\mathrm{DRR} = 10\log_{10}\frac{\int_{0}^{t_d} h^2(t)}{\int_{t_d}^{\infty} h^2(t)}, \qquad 1.6$$

where t_d is the direct-path propagation delay. In practice, the numerator of the equation 1. 6 is often assumed to be the largest magnitude peak at the beginning of the impulse response. Due to the finite sampling of RIR, the arrival time of the direct path component, and therefore the related energy, cannot be determined precisely. Therefore, a correction factor is often taken, which widen the time of the direct estimation path for a couple of ms.
DRR is inversely related to the distance of the sound source. This relationship results primarily from the effect of the inverse-square law on the direct portion of the sound field. Energy in the later-arriving reflected portion of the sound field may often be well approximated by a diffuse sound field, which is defined to have uniform energy over varying source positions. Hence, close sources produce a greater proportion of direct-path energy relative to the amount of reverberant energy than sources located farther away [28].

1.3.4 Early to Late Reverberation Ratio

The Early to Late Reverberation Ratio (ETL) gives the relative amount of early and late energy expressed in decibels. It is defined as:

$$\mathrm{ETL} = 10\log_{10}\frac{\int_{0}^{t_{mix}} h^2(t)}{\int_{t_{mix}}^{\infty} h^2(t)}, \qquad 1.7$$

where t_{mix} is the mixing time. When t_{mix} value is set to 50 or 80 ms, ELT is usually called Clarity Index. For car cabin application, the value of t_{mix} is of course smaller. A deeper investigation on t_{mix} in car cabin acoustic will be given in in Chapter 5.

1.4 Car wind noise

Wind noise inside a car cabin results from air flowing around the car body and increases with vehicle speed. At speeds above 100 km/h, it represents the major vehicle interior noise and it is critical in order to meet high quality customers expectations. In addition to airflow

excitation, vehicle interior wind noise perception depends on the sound transmission behaviour of the car body and may be influenced e.g. by the door and window seals [29]. Wind noise refers to the following noise and conditions [30] :

- Aerodynamic noise induced by the vehicle as it moves at high speed through a steady medium (air). This is related to the aerodynamic (or drag) coefficient of the vehicle, which is a function of the vehicle shape and of its cross-sectional area;
- Aerodynamic noise due to turbulence through “holes”, which corresponds to sealing to the vehicle parts;
- Aerodynamic noise due to exterior varying wind conditions, such as cross-wind on a highway. This is different from the previous two, since this type of wind noise is fluctuating.

In this thesis sources originating by the first group, i.e. caused by car’s high speed through a steady medium. This group is the most important and cannot be eliminated. The second general group occurs in “unhealthy” cars, due to errors in production or caused by damaged of sealing materials during use. The last group exhibits unsteady behaviour and produces rather low frequency noise (up to 300Hz), below Schroeder’s frequency, which is out of scope of this thesis.

In-vehicle wind noise can be localized exploiting, for example, beamforming methods. A schematic presentation of wind noise sources beamforming map is illustrated in Fig. 1-4. Based on this information, in-vehicle wind noise sources can be divided into 5 components. The first two correlated noise sources are caused by door side mirrors. They have a major impact on wind noise observed at the driver's position. Door mirrors are the largest outstands form a vehicle body, causing strong air turbulences behind them. Moreover, Left mirror is located very close to the drivers position.

The interaction between outside flows and exterior surfaces at the front and sides of the vehicle forms a strong swirling fluid structure called A-pillar vortex which is also an important wind noise sources [31]. Like door side mirror noise, it occurs symmetrically.

The last source that can be localized in car cabins is noise generated by wipers. The size of this source depends on the wipers size, but it is typically a distributed noise source located parallel to the bottom part of a windshield.

Fig. 1-4 Schematic presentation of a wind noise source beamforming map

Chapter 2

Head Related Transfer Functions for in-vehicle auralization

Summary

Quality of sound auralization highly depends on Head-Related Transfer Functions (HRTFs) used. Pinna, head and torso, which are considered in HRTFs, influences human aural perception in common daily life environments. However, when dealing with complex acoustic environments such as car cabins, higher accuracy of reproduced sound can be achieved if adding further body contributions in a HRTFs database. A new approach for measuring and processing HRTFs targeted at auralizing sound in car cabins has been proposed. The new HRTFs database includes reflections from almost straighten arms, like in a typical driving position, to improve sensation of externalization and enhance sound localization accuracy. Since some sources inside a car appear closer to driver's or passenger's ears (e.g. left mirror noise) and other farther (e.g. right noise), HRTFs are measured at several source- receiver distances. In order to improve repeatability of the approach and to have an acoustic input reference, volume velocity source is used as exciting source to measure HRTFs.

2.1 General overview of the HRTF database

Sounds can be reproduced in several ways, depending on the applications and requirements set. For some, a monaural signal can be sufficient. However, for sources where spatial information are important for correct sound evaluation binaural signals are required. In this case the simplest way is to record the signal using Head and Torso Simulator (HATS). In many situations, HATS recordings are not possible, therefore as an alternative very often a binaural synthesis. The synthesis is performed by convolving a source signal with a pair of Head-Related Transfer Functions (HRTF) filters in the time domain. HRTFs are essentially two transfer functions for each direction of sound incidence for the left and the right ear.

These transfer functions describe direction-dependent acoustic transformations of a sound from the free field, modified by the torso, head, pinna and ear canal before it hits the listener's eardrums.
HRTF databases are widely used in auralization. Some of databases are publically available. Recently a file format SOFA (Spatially Oriented Format for Acoustics) for storing HRTFs has been introduced [32] [33] and standardized by the Audio Engineering Society as AES69-2015. The most popular HRTF database referred in many publications are of the KEMAR dummy head [34] or Neuman KU 100 [35]. Well know and publically available are HRTF databases of humans measured at the entrance of an ear canal.
HRTF databases mentioned above are measured in far field conditions (above 1m source-receiver distance). Considering wind noise sources, this requirement is only fulfilled for a right mirror and an A-pillar. For closer wind noise sources (wipers, left mirror and A-pillar) a near field HRTF database should be used. This type of HRTF databases are also publically available are.
HRTFs describe how an impulse signal emitted by a point source in free field is affected by reflection from the torso, head, pinna and ear canal. However, driver's position is different from position of a listener's in a concert hall. Almost straighten arms creates the closest reflecting surface for acoustic waves, similar to the influence of the torso. None of know HRTF database include arms influence on HRTFs. Therefore, this chapter presents effect of typical driver's position, including straighten arms, on HRTFs.

2.2 Measurements and results

HRTFs are usually computed as the ratio between the complex sound pressure measured at the left and the right eardrums and the complex sound pressure at the center of the head with the head absent. By doing this division and assuming that exactly the same system was used to do both sound-pressure measurements, the transfer function of the measuring system which generally constitutes microphones, amplifiers, analog-to-digital and digital-to-analog converters, and loudspeakers is eliminated [36]. In this thesis, HRTF were measured using the LMS volume velocity source. The reason is twofold. This type of source incorporates an internal sound source strength sensor, which outputs a real-time volume velocity signal. Therefore there is no need to measure the pressure at the center of the head with the head absent, which increases measurement accuracy. The second reason for using this type of source is it's small dimensions. The source has small, 3cm width nozzle, which is important for near filed measurements.
HRTFs were measured in semi anechoic chamber. A Bruel&Kjaer 4128 HATS was used. Arms from a mannequin were attached to the HATS torso. In order to minimize the influence of different acoustic impedance of the HATS and the arms, a sweater was put on the modified

HATS as shown in the Fig. 2-1. During measurements, the excitation signal was a burst random noise of frequency range 200Hz to 20kHz. HRTFs were acquired using Test.Lab software and LMS Scadas frontend. Measurements were carried out in two session, first preliminary with azimuthal resolution of 30° and high arms inclination Fig. 2-1, and second more detailed with azimuthal resolution of 5° with lower arms inclination angle Fig. 2-2.

Table 2-1 HRTF measurement scenarios

Distance:	40 cm 70 cm 120 cm
Azimuth resolution	30 degrees 5 degrees
Elevation resolution	30 degrees
Arms inclination angles	2

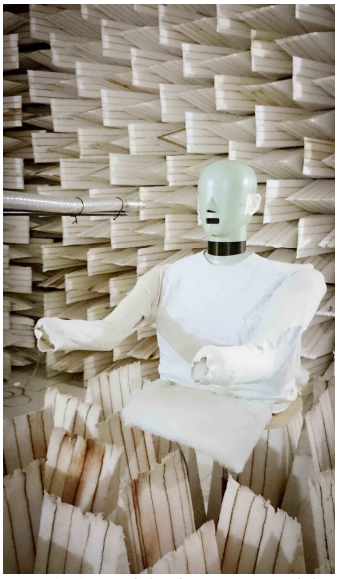


Fig. 2-1 Modified HATS used for measurement

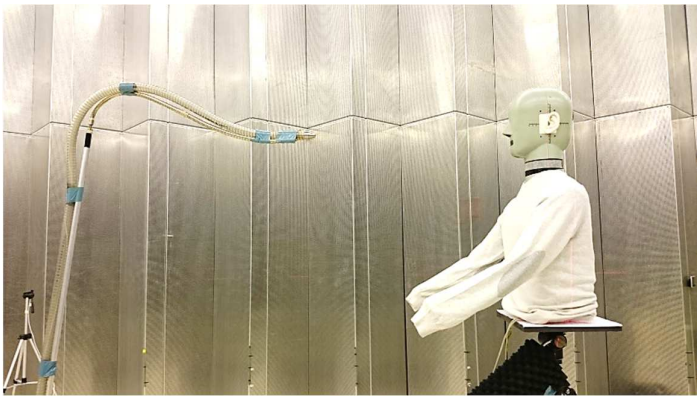


Fig. 2-2 Modified HATS used for measurement with lower arms inclination

During measurement campaign, HRTFs were acquired for different scenarios listed in Table 2-1 and Fig. 2-4 present HRTF and corresponding impulse response for 60^0 azimuth, 0^0 elevation and 70 cm distance. For the ipsilateral ear (the right ear in this case), influence of arms is very weak, whereas for the contralateral ear (the left ear in this case) HRFT is strongly affected by the reflections from arms. Contralateral ear is always in acoustic shadow, therefore reflections from arms have higher impact on the HRTF.

Reflections from arms for the contralateral ear create repeatable notches in HRTs. First notch can be observed at 1kHz and repeats every 1kHz. This difference in HRTF shape can cause a tone coloration.

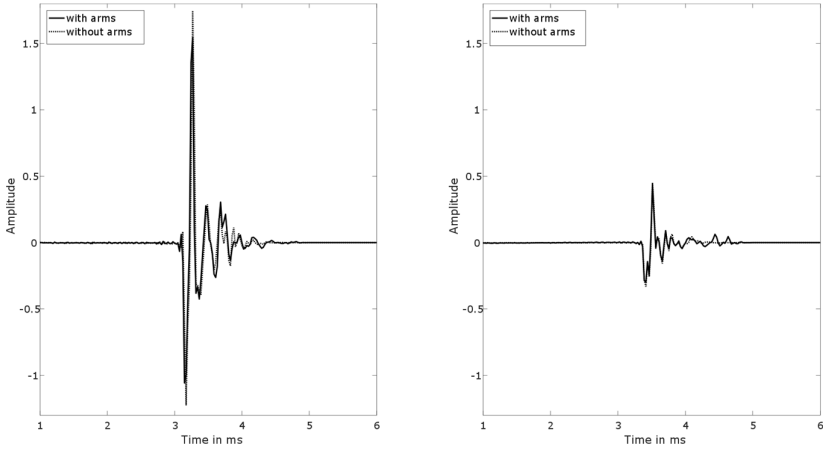


Fig. 2-3 HRIR with and without arms for 60º azimuth, 0º elevation and 70 cm distance.

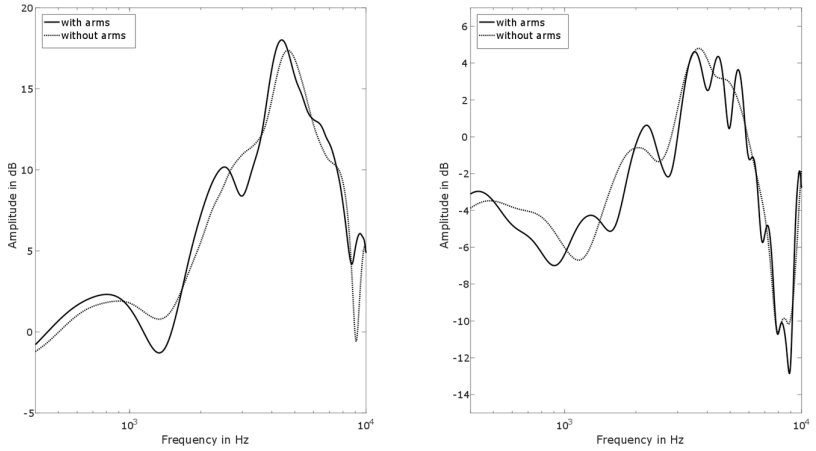


Fig. 2-4 HRTF with and without arms for 60º azimuth, 0º elevation and 70 cm distance.

Fig. 2-6 and Fig 2-6 show arms influence on near field HRTFs with respect to source-receiver distance for the contralateral ear. Presence of arms only slightly affects near field HRTFs. On the contrary, as source gets farther to receiver (more than 40cm), HRTFs have been found to change with distance due to presence of arms. At 70cm and 120cm distance reflections from arms are clearly noticeable, whereas they are less evident at 40cm. Such phenomenon is interesting for in-vehicle sound auralization since different source-receiver distances can occur. For example, in order to auralize left mirror and A-pillar wind noise, near field HRTFs should without arms can be use. For wipers, mirror and A-pillar on the right side of a car cabin, using the HRTF with arms database for is more suitable.

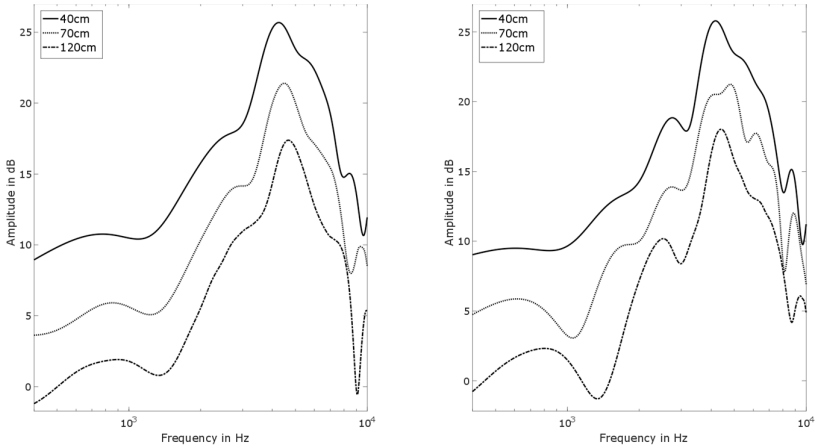


Fig. 2-5 Distance dependency of HRTFs for azimuths angles: 60˚(left) and 0˚ (right)

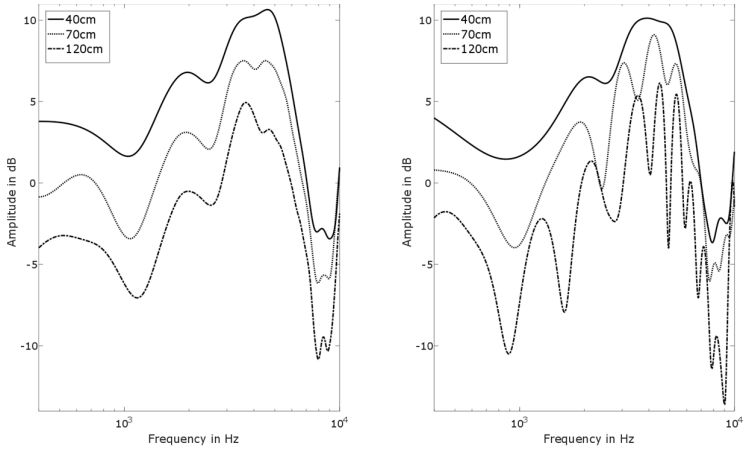


Fig. 2-6 Distance dependency of HRTFs for azimuths angles: 60˚(left) and 0˚ (right)

Fig. 2-8 and Fig. 2-8 show HRTF difference with arms and without arms for four azimuths angles: 0˚, 30˚, 60˚ and 90˚ for the ipsilateral and contralateral ear respectively at the distance of 70 cm and 120 cm. It can be seen that for 0˚, 30˚ azimuth influence of arms on HRTFs are approximately the same for both ears. For 60˚ and 90˚ azimuth angles for high frequencies differences between the contralateral ear and the ipsilateral ear becomes evident.

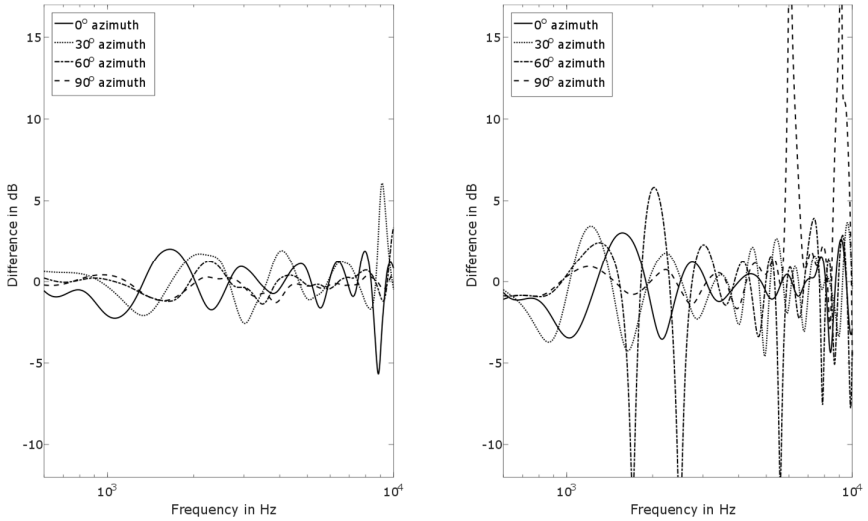


Fig. 2-7 Arms influence on HRTFs at the distance of 70 cm (left and right ear)

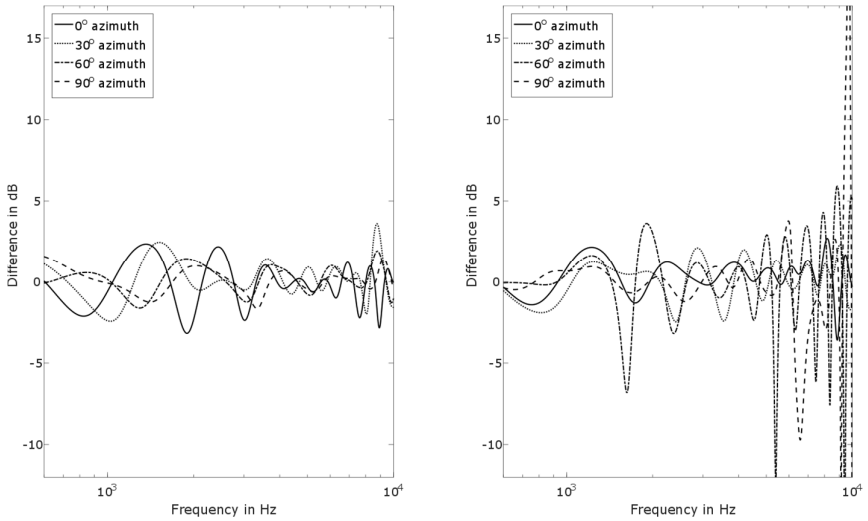


Fig. 2-8 Arms influence on HRTFs at the distance of 120 cm (left and right ear)

Fig. 2-9, Fig. 2-10 and Fig. 2-11 present HRTFs as functions of azimuth for 0° elevation at three different distances 120, 70 and 40 cm. These HRTFs were measured for HATS with lower arms inclination angle, as presented in Fig. 2-2. According to expectations, the influence of arms in HRTF in weaker for lower arms inclination angle. Nevertheless, differences are still clearly visible. It can be noticed that for the ipsilateral ear (azimuth from 0° to 180°) additional reflections from arms have the strongest impact on frequencies above 5-6kHz. For the contralateral ear (azimuth from 180° to 360°) a 'new' distinct pattern is visible for all frequencies.

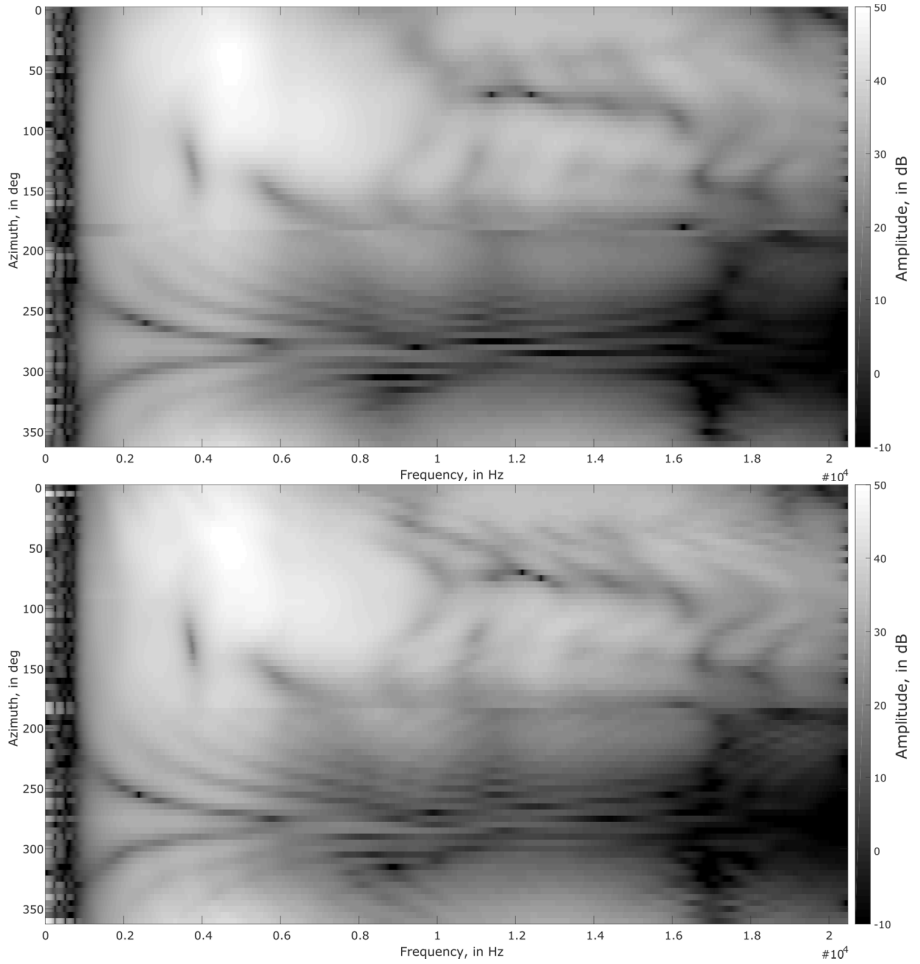


Fig. 2-9 HRTFs for different azimuth for 0° elevation at 120 cm distance without arms (top) and with arms (bottom)

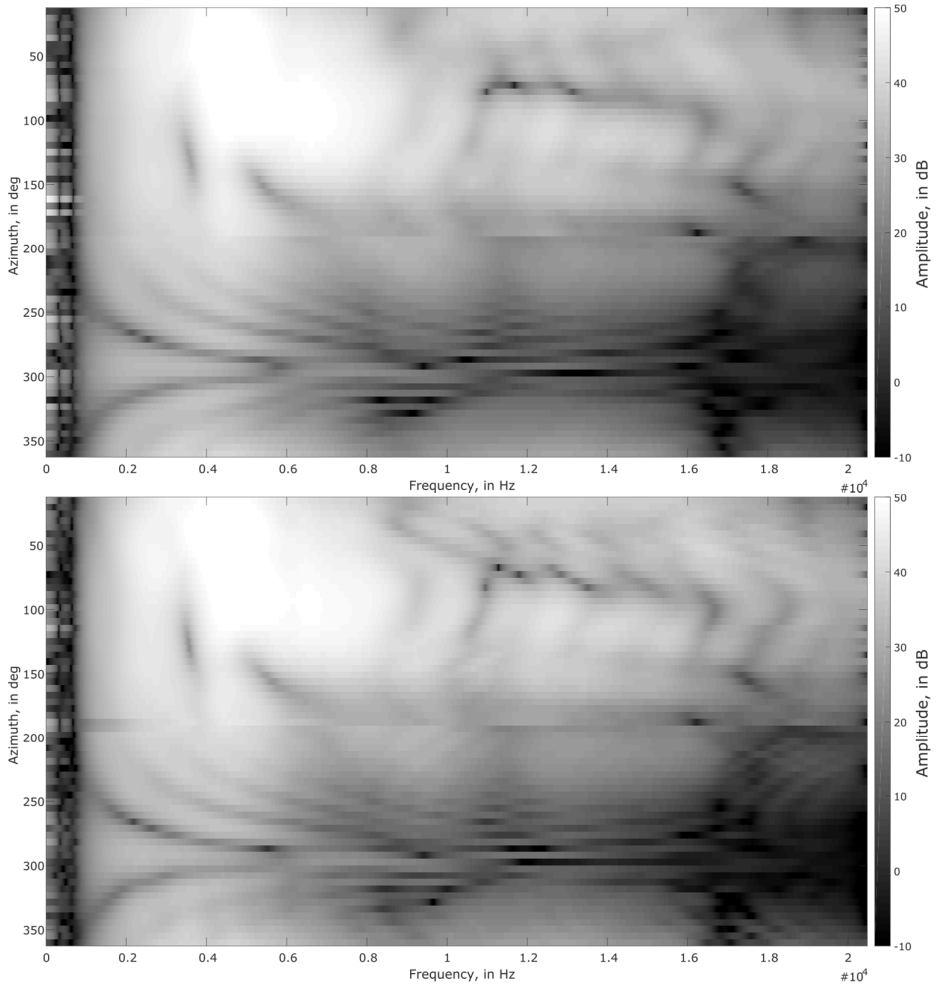


Fig. 2-10 HRTFs for different azimuth for 0 ^{0} elevation at 70 cm distance without arms (top) and with arms (bottom)

For 40 cm source-receiver distance, the influence of arms is very weak Fig. 2-11. The same effect was noticed for higher arms inclination angles Fig. 2-6. 40 cm distance is counted from the source nozzle to the center of the HATS's head, thus the source nozzle is really close to the HATS microphones. Acoustic waves reflected from arms reach HATS's ears mostly by diffraction.

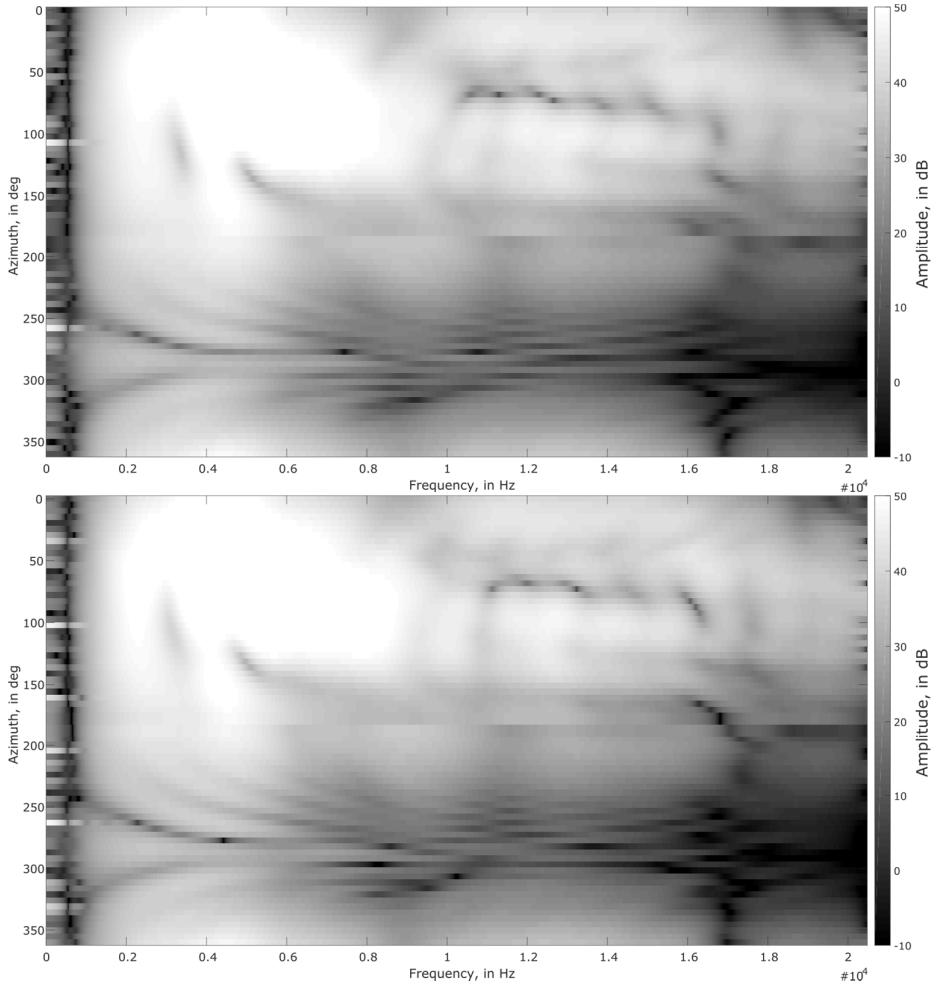


Fig. 2-11 HRTFs for different azimuth for 0^{0} elevation at 40 cm distance without arms (top) and with arms (bottom)

2.3 Conclusions

In this chapter, a new database for in-vehicle auralization war presented. Results show that presence of arms changes HRTF shape. Reflections from arms have the highest influence on HRTFs for azimuth higher than 20°. (when one ear is in acoustic shadow). Reflections from arms for the contralateral ear create repeatable notches in HRTs. First notch can be observed at 1kHz and repeats every 1kHz. This difference in HRTF shape can cause a tone coloration. Influence of arms is depend on distance. For 40 cm source-receiver distance, the influence of arms is very weak. For that close distance, acoustic waves reflected from arms reach HATS's ears only by diffraction. Therefore, HRTFs with arms should be used for wind noise sources appearing farther than 40 cm from the listener, e.g. wiper noise, right mirror and A-pillar noise. Despite this final conclusion, in the following chapters HRTFs without arms were used. This choice was made only due to required high spatial resolution of a HRTF database (1° instead of 5°).

Chapter 3

Auditory perception of early reflections in car cabins

Summary

A current challenge that NVH engineers are asked to face is to optimize the passengers' acoustic experience in vehicle cabins. The combined use of sound source localization techniques and auralization tools can help in tackling such issue. Data from the source localization step, i.e. sources and early reflections, can be auralized in order to assess the acoustic quality of the cabin. The remaining question is to understand how many reflections are effectively needed for providing a realistic listening experience. By varying the order of early reflections included in the auralization step, a different auditory impression is perceived by humans. At a certain point, increasing the number of individual reflections does not change the auditory impression and the rest of reverberations can be replaced by a constant reverberation tail. This chapter aims at defining the optimal number of reflections that is necessary to use for in-vehicle sound auralization purposes. A ray tracing model of a car cabin is used for generating the data set containing the sound samples auralized by changing the number of reflection orders. Results from subjective tests are presented to show the correlation between auditory impression in the cabin and the optimal number of reflections used in the auralization.

3.1 Early reflections problem in car-cabins

Auditory thresholds for early reflections in concert halls and rooms have been previously investigated using real sound sources [37, 38, 39] or using auralization for simulating direct and reflected sources [40, 41]. Additionally, some authors focused on minimum audible difference in direct to reverberant energy ratio [42, 43]. However, no previous work focusing on similar problem for automotive applications has been reported. Car cabins are characterized by peculiar acoustic environment, where reverberation time is short and

reflections appear close to each other in time. Different, compared to concert halls or rooms, forward and backward masking [5] can cause different perception of early reflections.
In this chapter two problems are investigated. First, how including or excluding individual reflections in auralization affect human perception (threshold for inclusion of a reflection). Car interior is a mixture of various absorbing materials (e.g. well dumped seats, doors and reflective windows), where we can expect single strong reflections from windshield or windows, and many weak reflections from trimmed parts. Hence, only a few early reflections might be necessary to include in auralization to achieve the "natural" sound, while the rest of reverberations can be replaced by a stochastic reverberation tail as will be described later in Chapter 5. From Chapter 1 we know that main wind noise contributors are mirrors, A-pillars and wipers. The main focus of this chapter is on a left mirror, as the most dominant source and closest to a drivers' position. Analysis was carried out on left mirror noise separately and together with other wind noise sources as markers.
The second problem under investigation is the level of early reflections at which we can perceive change in auralized sound (threshold for inclusion of an early reflection). This threshold leads to very useful information of required dynamic range of sound source localization techniques for car cabins applications. Previous studies show threshold dependence as a function of angle of incidence, time delay and sound source type [3, 44, 45, 42, 46]. Therefore, results vary from almost -30 dB (reflection to direct relative level) for broadband noise, lateral reflection and long delay time [39], to -5 dB for speech, reflection from the same direction and short time delay [41]. Also dependence of the threshold level and reverberant environment was reported [39]. This chapter derives threshold of individual reflections for short delay times, broad band noise simulating car cabin wind noise and different directions (detailed description in section 3.1). As in the problem of threshold for inclusion of a reflection, analysis was carried out on left mirror noise separately and together with other wind noise sources as markers.

3.2 Experiment investigation

In order to determine auditory thresholds of early reflections in car cabins, four different scenarios were investigated:

- Only left mirror wind noise source with removed reflections;
- Left mirror wind noise source with removed reflections in the presence of other wind noise sources;
- Only left mirror wind noise source with amplified or deamplified reflections;
- Left mirror wind noise source with amplified or deamplified reflections in the presence of other wind noise sources.

Using ray tracing technique, impulse responses of a premium sedan car cabin were generated. Calculations were carried out in LMS Virtual.Lab software. In order to simulate wind noise, a monopole source was placed on the left mirror. The dry sound was obtained by filtering a white noise to achieve the same noise spectrum as measured in a wind tunnel during previous measurements. Fig. 3-1 reports the wind noise spectrum used for auralization. Receiver was placed in the center of a driver's head. A model of the car with marked ray paths and the corresponding energetic impulse response are presented in Fig. 3-2.

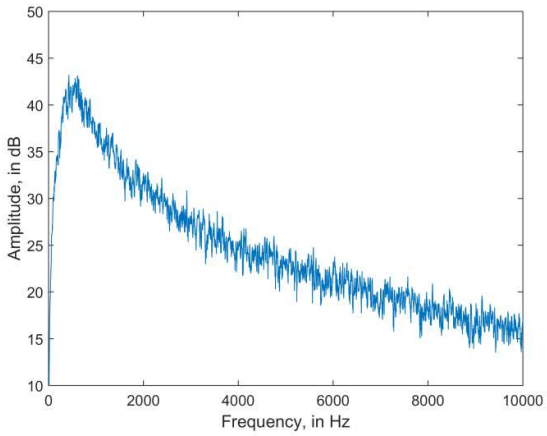


Fig. 3-1 Spectrum of wind noise signal used in auralization.

In practice, not only one mirror generates noise. As illustrated in Fig. 1-4, both mirrors, A-pillars and wipers are acting together. Thus, it is reasonable to investigate our problem in the presence of all sources. For simplification, only left mirror noise will be modified, whereas all other sources play the role of a "masker". The principle of our "masker" calculation stays the same. In the car model, sources were placed in positions corresponding to natural occurrence, receiver in the centre of a drivers' head and the ray tracing technique was used to calculate impulse responses. Detailed description of source location in the car cabin and levels are presented in Fig. 3-2

The ray tracing model was used not only to generate impulse response of a car, but also to obtain direction of arrival of each reflection, which is a great advantage of ray tracing over other numerical methods. The reverberant stimuli were generated by additional processing of the dry stimuli. Each reflection was spatialized by convolution with HRTF filter pair, corresponding to the angle of incidence calculated by ray tracing (see [47] for detailed description of HRTF database used).

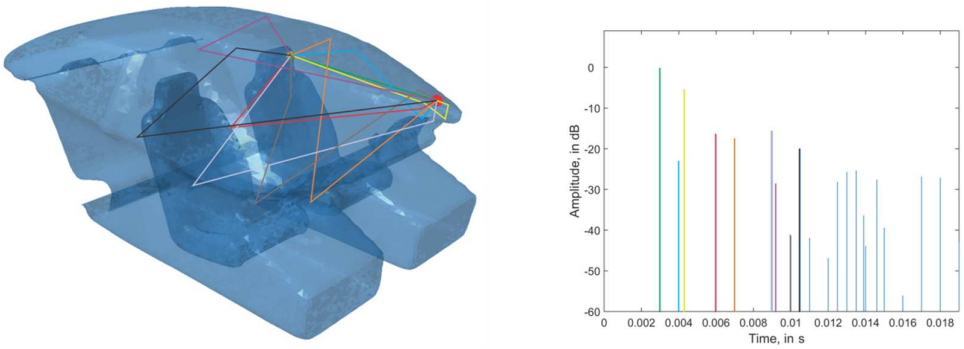


Fig. 3-2 The model of a car used in ray tracing simulations (left) and first 20ms of the corresponding impulse response (right).

Four listening experiments were carried out, one for each scenario. The aim of the first and the second experiment was to determine the threshold for inclusion of a reflection (early or late). The last two experiments were focused on the level of early reflections at which humans can perceive change in auralized sound. Ten listeners without known hearing problems (all male, ages 22–31) participated in all experiments. All stimuli were presented to subjects via headphones (Sennheiser HD-600 with an external sound card) within a quiet office room, where the noise floor was below 20 dBA. The level of stimuli was calibrated using a head and torso simulator in order to present them with a level comparable to the one generated by wind in car cabins – 45 dBA.

In the first experiment, sounds, which were presented in pairs, consisted in a reference reverberant sound (spatialized, with all reflections) and a modified sound. The modifications consisted in excluding individual or multiple reflections from the impulse response. Finally, listeners were asked to identify whether two stimuli heard in succession are different or not. The test consisted of 50 pairs, and reference and signals were presented in random order, with an interstimulus interval of 500 ms. Due to the high number of reflections in echogram, performing full factorial experiment is practically impossible. Therefore, preliminary investigations took place, in order to select the most relevant reflections or group of reflections to be removed in stimuli for the experiment. Reflections from one to eight (those appearing within first 10ms, see Fig. 3-2) were removed separately at a time. Most of sound samples consisted of two or three removed early reflections. The purpose was to investigate the influence of each early reflection separately and a group of early reflections on auditory perception. The latter, gives as information about possible interactions between early reflections. Last part of stimuli was focused on a late reverberation. Thus, we removed some of reflections appearing after 10ms. Detailed description of which reflections were removed from sounds and corresponding results are listed in Table 3-1.

The second experiment is essentially similar to the first one. The only difference is an added masker consisting of other wind noise sources.
The third experiment was carried out in a similar manner as the first one. However, the way how stimuli were modified is different. Here, we were interested in the level of early reflections at which humans can perceive change in auralized sound. Therefore, a set of stimuli were prepared where the level of individual reflections was changed, while other reflections remain the same. The level of analyzed reflection was changed in 1dB step from -18dB to -8dB (reflection to direct relative level). Due to time constrains of a listening experiment and preliminary investigations, only the first five reflections were taken into consideration. To sum up, 55 pairs, consisted of modified sound and the reference, were presented to listeners. Like in previous experiment reference and signals were presented in random order, with a time gap of 500ms and listeners were asked to identify whether two stimuli heard in succession are different or not. The last experiment repeats experiment no. 3 with added noise, like in experiment no. 2

3.3 Results and discussion

In all conducted experiments, thresholds were calculated analogously. Each modified sound gets 1 point when a listener identifies it as different from the reference. Points for all stimuli for all participants are summed up together and divided by the number of participants. In this way, results are presented on a scale from 0 to 1, where 0 means none of participants was able to recognize a difference, and 1 means all participants perceived the stimulus as different from the reference. The threshold is a point where at least half of participants identified a difference (results higher or equal 0.5).

3.3.1 Threshold for inclusion of a reflection

Table 3-1 presents results of the experiment for the left mirror; reflections removed from stimuli are listed as well in Table 3-1.
analogously presents results for the left mirror in the presence of all wind noise sources. According to predictions, reflection number two was selected to be the most significant in both experiments. Lack of this reflection, independently or together with other reflections, is audible by all listeners who participated in the experiment. The reason is straightforward: the level of this reflection is approximately only 5 dB lower than the true source which is in compliance with results achieved in the next section. Additionally, reflection number two appears only 1,3ms after the true source. That short repetition of a “dry” sound coming from almost the same direction as the source causes strong interference pattern. This leads to a perceptual effect called tone coloration [3, 45]. Fig. 3-3 shows two spectra of the left mirror

Table 3-1 List of sound samples used in the experiment with numbers of removed reflections and corresponding results

Sample no.	Reflections removed	Score for left mirror	Sample no.	Reflections removed	Score for left mirror
1	1,2	1	26	1,8	0,3
2	2,3,4,5	1	27	3,7	0,3
3	2,3,4	1	28	5	0,3
4	2,3	1	29	1,3	0,2
5	2,4	1	30	1,7	0,2
6	2,5	1	31	1	0,2
7	2,6	1	32	3,6	0,2
8	2,8	1	33	3	0,2
9	2	1	34	4,6	0,2
10	3,5	1	35	4	0,2
11	2,7	0,9	36	5,6,7	0,2
12	3,4,5	0,8	37	5,6	0,2
13	1,3,4	0,7	38	6,7,8	0,2
14	4,9	0,7	39	6	0,2
15	4,5,6	0,6	40	7	0,2
16	4,5	0,6	41	8	0,2
17	5,6,7,8	0,6	42	13,14,15,16	0,2
18	5,7	0,6	43	4,8	0,1
19	1,3,4,6,7,8,9	0,5	44	6,7,8,9,10	0,1
20	3,4	0,5	45	6,8	0,1
21	4,5,6,7	0,5	46	7,8	0,1
22	4,7	0,4	47	13,15,18,21,23, 24,25	0,1
23	5,8	0,4	48	1,6	0
24	1,4	0,3	49	6,7	0
25	1,5	0,3	50	9,10,11,12	0

Table 3-2 List of sound samples used in the experiment with numbers of removed reflections and corresponding results for scenarios with

Sample no.	Reflections removed	Score for left mirror with noise	Sample no.	Reflections removed	Score for left mirror with noise
1	1,2	1	26	13,14,15,16	0,2
2	2,3,4,5	1	27	6,7,8,9,10	0,2
3	2,3,4	1	28	3,4,5	0,1
4	2,3	1	29	5,6,7,8	0,1
5	2,4	1	30	5,7	0,1
6	2,5	1	31	3,4	0,1
7	2,6	1	32	4,5,6,7	0,1
8	2,8	1	33	5,8	0,1
9	2	1	34	1,5	0,1
10	2,7	1	35	3,7	0,1
11	1,3,4	0,3	36	5	0,1
12	1,7	0,3	37	1,3	0,1
13	3,5	0,2	38	3,6	0,1
14	4,9	0,2	39	4,8	0,1
15	4,5,6	0,2	40	6,8	0,1
16	4,5	0,2	41	7,8	0,1
17	1,3,4,6,7,8,9	0,2	42	13,15,18,21,23, 24,25	0,1
18	4,7	0,2	43	1,4	0
19	1	0,2	44	1,8	0
20	4,6	0,2	45	3	0
21	5,6,7	0,2	46	4	0
22	5,6	0,2	47	6,7,8	0
23	6	0,2	48	1,6	0
24	7	0,2	49	6,7	0
25	8	0,2	50	9,10,11,12	0

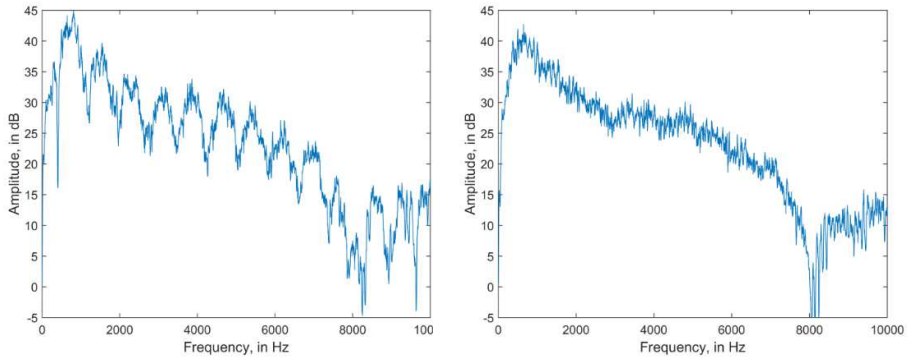


Fig. 3-3 Spectrum of the left mirror wind noise with all reflections (left) and sound without first five reflections (right)

noise with all reflections (left) and the same noise without first five reflections (right). Strong comb filter can be noticed when all reflections are present. Removing early reflections makes the spectrum more "flat", therefore reduce tone coloration. Further removing of reflections does not affect a change in the spectrum and becomes unnoticeable for listeners. The spectrum presented in left part of Fig. 3-3 differs from the "dry" sound (Fig. 3-1) because the same dry sound was convolved with the HRTF filter related to the direct path.

Removing reflections from the late reverberation (after 10ms) does not affect perception of a sound. As can be seen from Table 3-1, stimuli no 42, 44, 47 and 50, where a high number of reflections was removed, were impossible to distinguish from the reference stimulus. This leads to a conclusion that reflections from late reverberations do not need to be treated individually. It is sufficient, for an auralization, to replace the exact calculation of a late reverberation, e.g. from a ray tracing simulation, with an universal reverberation tail, calculated stochastically.

Table 3-3 Levels of first eight reflections with corresponding directions of arrival and an indications of significance

Reflections	1	2	3	4	5	6	7	8
Level, in dB	-23	-5	-16	-17	-15	-28	-41	-20
Azimuth, in deg	340	345	90	280	75	215	45	80
Elevation, in deg	10	10	-40	0	-20	20	10	0
Significance	no	yes	yes/no	yes/no	yes/no	no	no	yes/no

3.3.2 Threshold for an individual reflection with a full impulse response

Fig. 3-4 and Fig. 3-5 present results from experiments where perception thresholds for individual reflections were investigated in two scenarios: with only one noise source acting and with all sources acting together. Results are presented in reflection to direct relative level. Fig. 3-4 shows results of the third experiment. Reflections 2, 3 and 5 have lower threshold (around -14 dB, reflection to direct relative level) than reflections 1 and 4 (around -10 dB). Reflection no.1 appears almost together from almost the same direction with very strong reflection no.2, thus higher threshold can be explained by masking effect caused by the reflection no.2. Despite the fact that reflections 3 and 4 have almost the same levels they have different thresholds of perception. The reason can be explained by differences in direction of arrival. Reflection no. 3 is lateral, causing higher energy in the right ear than the reflection no. 2. The fact that reflections appear from opposite directions affects changes in frequency content of an auralized noise simply due to convolution with different HRTF filters.
When not only one source is present at the same time, detection of individual reflections is different. Fig. 3-5 presents results of the experiment when the RIR of left mirror noise source is modified and all wind noise sources i.e. from right mirror, A-pillars and wiper are added as a masker. In this scenario, perception thresholds of individual reflections are the same for all early reflections. Differences between thresholds are within 1dB which can be neglected. Comparing this results with the scenario without masker (Fig. 3-4) shows that all reflections have higher perception thresholds, except reflection no.4 where the threshold stays the same. Interesting is that when masker is added, perception threshold are independent from direction of arrival. Because masker signal consist of other wind noise sources practically from all direction, some spatial cues (like spectrum differences) are distorted, therefore thresholds depends on level differences only.

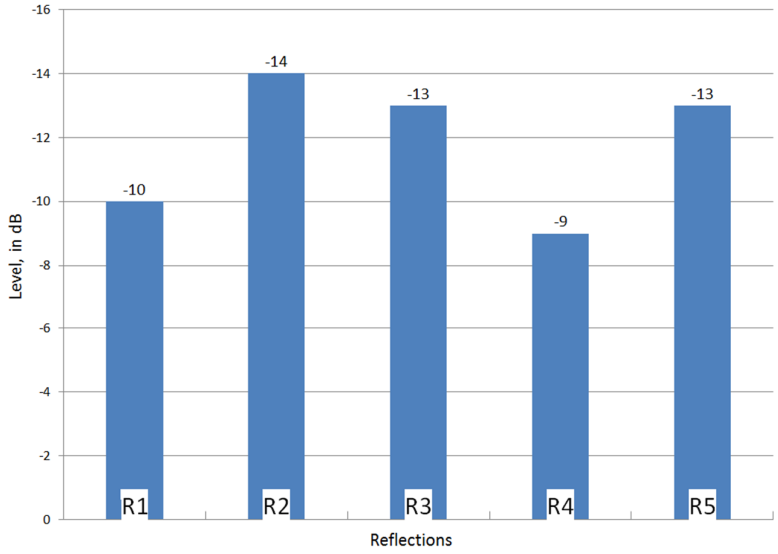


Fig. 3-4 Thresholds for single reflection with a full impulse response

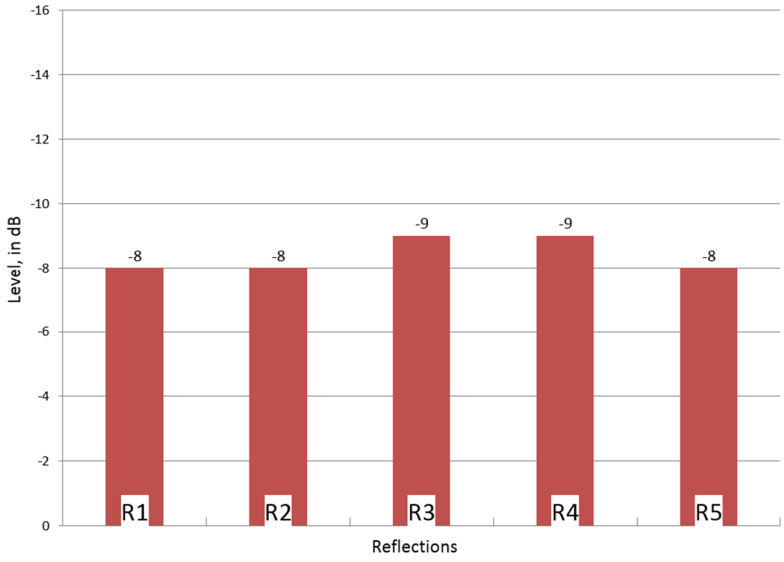


Fig. 3-5 Thresholds for single reflection with a full impulse response and all wind noise sources present

Listeners in this experiment were consistent. Influence of results where listeners did not perceive a difference of a stimulus with a level higher than beforehand marked as a perceived is statistically non-significant.

3.4 Conclusions

To conclude, it is possible to reproduce a sound in car-cabins by reducing number of reflections included in auralization. Considering wind noise source, optimal number of reflections used in the auralization is defined by the level of individual reflection as well as the direction of arrival. For reflections appearing from a similar direction as a direct sound, important are those with a level higher than -10dB relative to a direct sound. Listeners are more sensitive for lateral reflections, where the threshold is approximate -14dB relative to a direct sound. When all wind noise sources are acting together, perception of early reflections in car cabins decreases and it becomes independent form direction of arrival of the reflection.

Chapter 4

Sound source localization accuracy in car-cabins: a human perception perspective

Summary

The aim of this chapter is to present a human-oriented approach to define the minimum accuracy required for source localization techniques when used in car interiors in combination with auralization strategies. Two experiments were carried out based on simulated and measurement data. In the first experiment a ray-tracing model of a car cabin is used as data provider. By assigning a random distribution of the source positions around the ideal one, the associated reflections are then consequently estimated from a ray-tracing model of the cavity. Source and reflections signals are convolved with corresponding Head Related Transfer Functions (HRTFs) for generating data aimed at filling a data base for subjective evaluation. In the second experiment, measured impulse responses were taken and convolved with free-field wind noise signal. Listening experiments were conducted in order to understand the influence of both source localization accuracy and different reverberant environments (due to reflections) on human perception. A Minimum Audible Angle MAA is also proposed as a parameter to assess the Just Noticeable Difference (JND) for sound source localization measurements inside a car. The influence of multiple wind sources acting at the same tame is also presented.

4.1 MAA overview

Optimizing the passengers' acoustic experience in vehicle cabin is becoming the most important task in the NVH field. However, any improvement to the cabin acoustic response starts from the identification of the source locations entering the cabin. Among the different sound source localization techniques, acoustic beamforming has gained relevance in car interior applications, mainly because it guarantees a good balance between localization

accuracy and contained testing time. Moreover, interior beamforming can provide information related to the main source position as well to the early reflections. For these reasons, a trend in the recent years is to exploit data collected in a beamforming test also for auralization purposes. Despite this is an interesting approach, the fidelity of auralization strongly depends on the accuracy of localized sources and reflections. This issue has been faced, in the past, addressing energetic aspects only [48]. This chapter aims at identifying the accuracy required for a sound source localization technique, as acoustic beamforming, in identifying the position of sound sources in car cabins from a human-centred point of view, i.e. when the source location are to be used for auralization purposes. This is done, here, by associating localization accuracy to the Minimum Audible Angle (MAA), the angle formed at the centre of the head by lines projecting to two sources of sound whose positions are just noticeably different [49].

A sound source localized in a position that differs from its true location may affect perception of the environment during a sound reproduction. In free field, this difference is caused only by the change in direction of arrival of the localized source. For sinusoidal signals presented on the horizontal plane, spatial resolution is highest for sounds coming from the median plane (directly in front of the listener) with about 1° MAA, and it deteriorates markedly when stimuli are moved to the side e.g., the MAA is about 7° for sounds originating at 75° to the side [50, 51]. MAA depends not only on direction, but also on the type of signal and on the frequency content. The lowest values of MAA in the horizontal plane are around 750Hz (MAA=1°), and the highest around 2kHz (MAA=3°) [50].

Those differences have their origin in spatial cues which are provided to both ears. The sound reaching the farther ear is delayed in time and is less intense than that reaching the nearer ear. There are thus two possible cues as to the location of the sound source: an interaural time difference (ITD) and an interaural level difference (ILD) [51]. In sound propagation in reverberant environment, such as car cabins, the first event is the direct sound (the "first wave-front"). This is followed by a set of sparse early reflections for the next 20 ms or so (depending upon the size of the car), and then by dense late reflections that form a decaying "tail". Different position of a sound source affects then different energy of reflections and distribution over time. Thus, coming back to our original problem, inaccurate localization of source in a car cabin does not only leads to a wrong direction of arrival, but to different reverberation pattern as well. This might cause other perceptual effects like tone coloration or image shift [52]. Therefore, this chapter will focus on perceptual effects of different source positions investigating any perceptual differences, not only direction of arrival.

4.2 MAA for one source in car cabins

4.2.1 Stimulus generation

In this section investigations of MAA are based on a simulated scenario. A car mock-up model was created in LMS Virtual.Lab software in order to generate impulse responses. A ray tracing technique was employed for calculations. A simulated wind noise source was used for the experiment. The dry sound was obtained by filtering a white noise to achieve the same noise spectrum as the one measured in a wind tunnel during previous measurements. Fig. 3-1 reports the wind noise spectrum used for later auralization.
The receiver was placed in the position of a driver's centre of the head. Sources were distributed evenly on the windshield with 2° angle step within the range of 40°, where 0° indicates driver's centre of the head. All sounds were distributed either on the median plane (Fig. 4-1) or on the horizontal plane (Fig. 4-2).

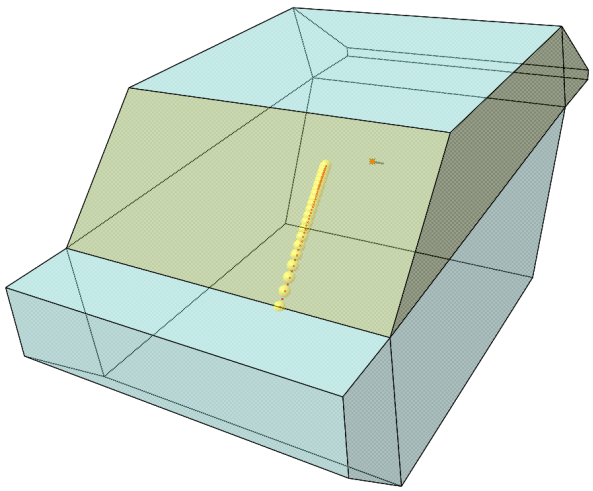


Fig. 4-1: A car mock-up model used for generating impulse responses with marked sources in the median plane.

The ray tracing model was used not only to generate impulse responses of the simplified car mock-up, but also to obtain direction of arrival of each reflection. The reverberant stimuli were generated by additional processing of the dry stimuli. Each reflection was spatialized by convolution with HRTF filter pair, corresponding to the angle of incidence calculated by ray tracing (see [35] for detailed description of HRTF database used).

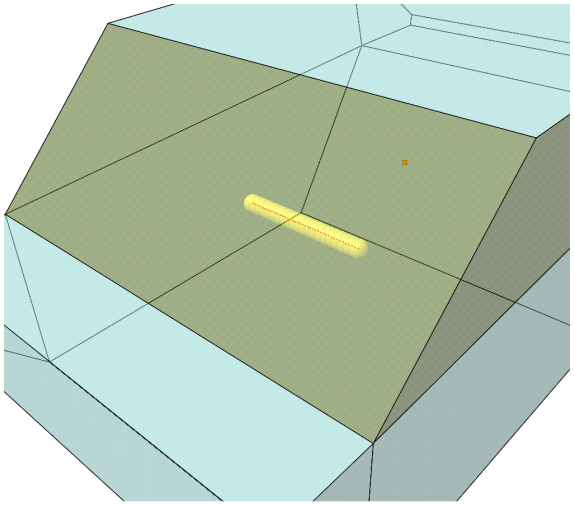


Fig. 4-2: A close-up of a car mock-up model used for generating impulse responses with marked sources in the horizontal plane.

4.2.2 Experiment investigation

In order to determine the MAA for the azimuth and elevation, a listening experiment was carried out. Eleven listeners without known hearing problems (two females and nine males, ages 22–35) participated in the experiment. All stimuli were presented to subjects via headphones (Sennheiser HD-580) in a quiet office room, where the noise floor was below 20 dBA. In the experiment, sounds were presented in pairs. Each pair consisted of a reference sound and a tested sound. The reference sound was a sound appearing in forward direction (0° elevation and 0° azimuth angle). Tested sounds varied in median and horizontal plane within the range of 40°, where 0° indicates driver's centre of the head. If the listener perceived a difference, another test sound, which is closer to the reference, was presented. It is assumed for higher angles to be easily detectable, thus tested sounds were presented in decreasing angle order, e.g. from 6° to 4°. MAA is defined as a minimum angle at which the listener can perceive a difference between the tested sound and the reference. At first, sounds related to different azimuth angle in the vertical plane were investigated. Then, sounds varied in median plane with constant 0° azimuthal angle were presented. The order of appearance of the reference sound and the sound under investigation was randomized.

4.2.3 Results and discussion

Fig. *4-3* presents results achieved from the listening experiment. MAA for the median plane was found to be equal 6° for upper position, and 4 ° for lower position of the source. For the

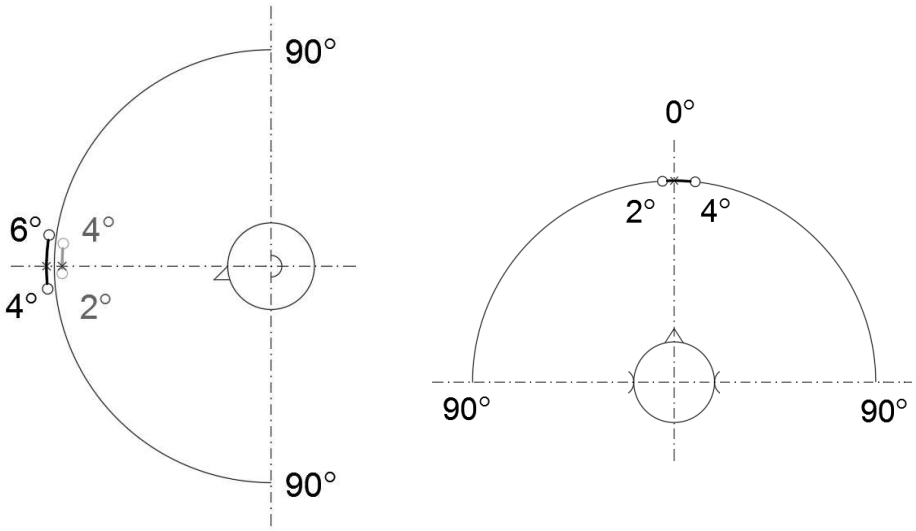


Fig. 4-3: MAA for the horizontal direction (right) and for the vertical direction (left). For MAA on the median plane (left) results are marked with black line and with grey line for 4° azimuth.

horizontal plane, MAA is different for left and right source positions, which is 2° and 4° respectively. Listeners responses are presented in Fig. 4-4.

Due to the fact that localization in the vertical direction varies with azimuth, an additional test was carried out. After estimation of MAA in the horizontal plane, sound samples were generated for 4° azimuth and different elevation angles. Results are presented in Fig. 4-4. In all tests, listeners responses varied within 2° only, which is equal to the experiment accuracy. In general, we need to remember that achieved results should not be directly associated with localization of a sound. Perceived difference between two sounds appearing from different direction corresponds to different location of the source only in free field conditions. In this chapter sounds inside a car cabin were investigated, where besides direct sound, multiple reflection occurs. Listeners were asked to assess the difference between two sounds, not a perceived direction of arrival. Nevertheless, achieved values of MAA are similar to those

known from the free field conditions. For MAA in median plane in car cabins is the same as for the white noise in free field. In the horizontal plane, on the contrary, MAA is slightly worse than in the free field. Inconsistency of the left and right MAA in horizontal plane and up and down for 4° azimuth vertical plane was also reported. Possible differences listed above might be due to following reasons:

a. Experiment accuracy is equal to 2°. The accuracy is limited to a resolution of the HRTF data base used. Moreover, 2° corresponds to a very small change of a source displacement: 1.5cm;

b. Even though in reverberant environments according to the precedence effect [2] sound source is perceived from the direction of the first wave front, in car cabins early reflections appear very close to each other in time. This leads to other perceptual effects like tone coloration or image shift [52, 45]. For some displacements distributions of early reflections may provide different audible effect, not directly correlated with sound localization;

c. Position of the source inside a car. For some sources positions, especially those close to boundaries, even small displacement may lead to perceptual difference. This case can be noticed in every estimated MAA. Moving source on the horizontal plane towards left is more likely to perceive than the same displacement towards right (Fig. 4-3 right). Source moving towards left is closer to the car's left door, and therefore reflection from this obstacle arrives faster than when the source is displaced towards right;

d. Source frequency content. In this experiment a wind noise signal was used, which is similar to Brown noise centred at 500 Hz. For more broadband signals, due to higher content of useful ILD and ITD, source displacements might be perceived even more accurately. There exists also the possibility that, for some signals, even more significant changes of source position might stay unnoticed.

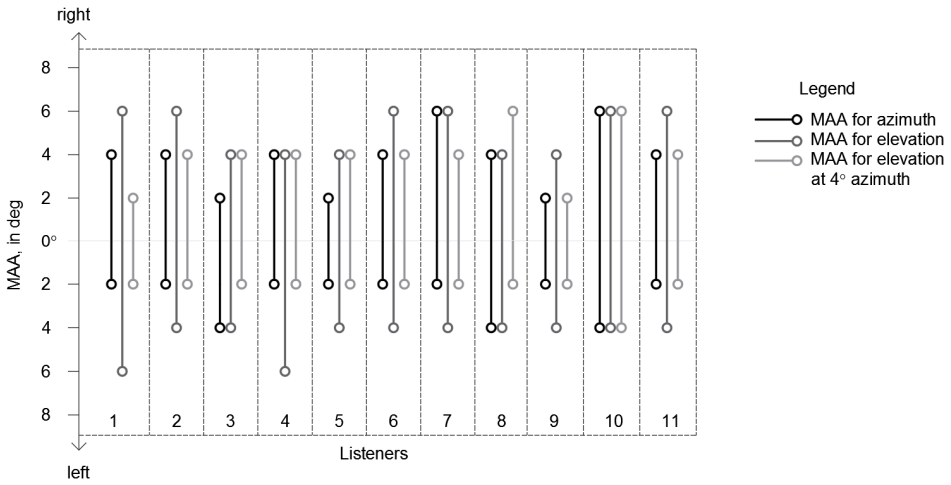


Fig. 4-4: Results of MAA for the horizontal plane, median plane and vertical plane at 4 ° for each listener.

MAA for known source-receiver distance can be represented in cm. For the investigated car cabin mock-up 2° corresponds to approximately 1.5 cm. To summarize, differences caused by source displacements larger than 3 cm were clearly heard by all participants. In some cases, e.g. displacements in the horizontal plane towards left were perceived with the change as small as 1.5 cm. Those results are in compliance with similar work [53], where small displacements of a source in car cabins were investigated based on binaural records.

4.3 MAA for multiple sources

In the Section 4.2 investigations pertained to one source acting at a time. Intuitively, MAA in this conditions seems to be the smallest. As stated in the previous section MAA, and as a result minimal audible displacements of a source, are really small. In a real situation, multiple wind noise sources are acting simultaneously. Therefore, it is natural to ask a question: how the value of MAA would change in presence of other signals? It can be supposed that, for instance, MAA for the left mirror, when also right mirror and wipers are acting together, will be higher than in a situation when only the left mirror is present. But is MAA the same for all wind noise sources? When localizing wind noise sources, which source should require the highest accuracy? This section will try to answer these questions.

To increase accuracy of MAA estimation, IRs of a car were measured instead of calculating them using a ray tracing model. In simulated scenario, accuracy of HRTFs database

resolution needed for sound spatialization and ray tracing calculation itself are sources of uncertainty. By measuring IRs using HATS we can obtain the most accurate results, unaffected by HRTF database resolution or interpolation (typically 2°). The acquisition was performed using an LMS SCADAS frontend together with LMS Test.Lab software. G.R.A.S. KEMAR HATS was used as a receiver. In this type of measurement it is crucial to keep the receiver in exactly the same place. Therefore position of the HATS was controlled by two laser beams pointed at the left ear. A new compact LMS Qsources volume source was utilized. Such a source incorporates an internal sound source strength sensor, which outputs a real-time volume acceleration signal, and it emits the noise as a monopole source from 1kHz to 10 kHz. Fig. 4-6 presents HATS and compact LMS Qsource used for IRs acquisition. Fig. Three noise sources were taken into account for investigations: left and right mirror noise and wipers noise. LMS Qsource was therefore placed in several uniformly distributed positions with respect to drivers' centre of the head. Fig. 4-5 presents measurement points on the left and right windows and on the windshield. For the left and right windows, measurement points positions were determined by the geometry of front door windows. A mirror noise can appear on a beamforming map everywhere around the front-bottom corner of a window, therefore IRs were measured as close to the window corners as possible. Wipers noise appear on beamforming map at the bottom of the windshield. To simplify later investigations, wipers noise was assumed to be a monopole point source at the middle of the bottom part of the windshield. Table 4-1, Table 4-2 and Table 4-3 present source shift values in degrees with respect to the reference position (0° for elevation and azimuth) for the left and right windows and the windshield, respectively.

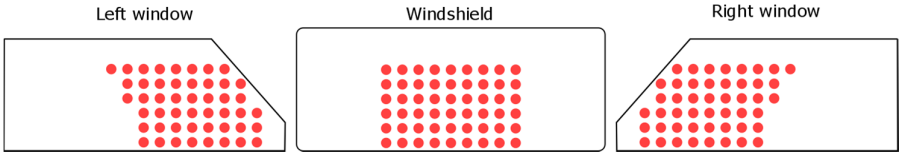


Fig. 4-5 Measurement points on the left and right windows and the windshield.

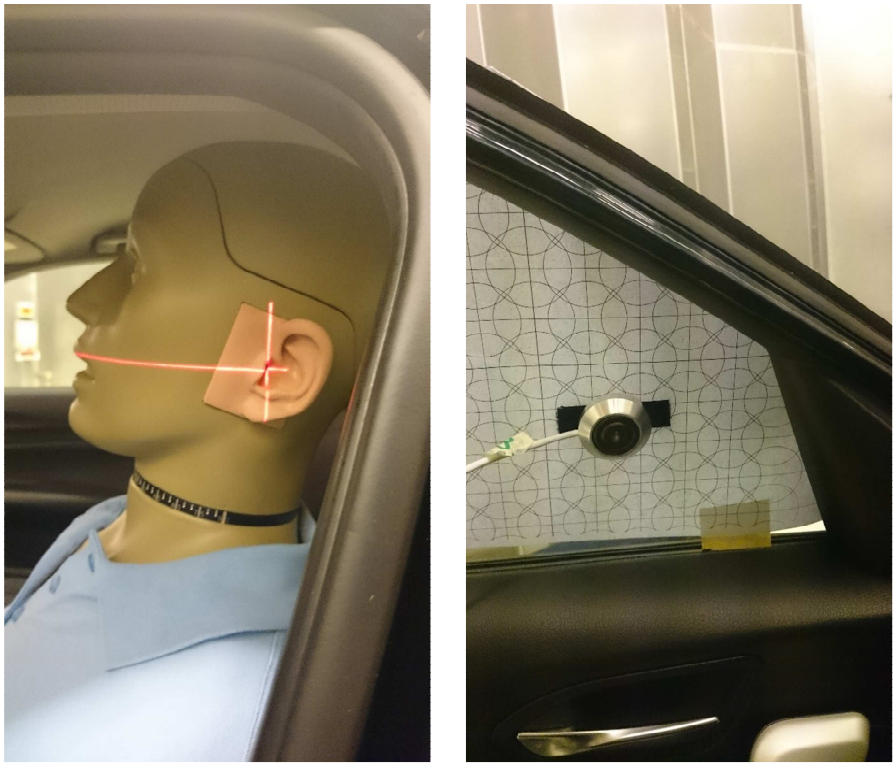


Fig. 4-6 KEMAR HATS used for IRs acquisition (left) and compact LMS Qsource at the left window (right).

Table 4-1 Source shift values in degrees with respect to the reference position for the left window.

ELEV	AZIM										
6	10	8	6	4	2	0	-2	-4	-6	-	-
4	-	8	6	4	2	0	-2	-4	-6	-8	-
2	-	8	6	4	2	0	-2	-4	-6	-8	-
0	-	-	6	4	2	0	-2	-4	-6	-8	-10
-2	-	-	6	4	2	0	-2	-4	-6	-8	-10
-4	-	-	6	4	2	0	-2	-4	-6	-8	-10

Table 4-2 Source shift values in degrees with respect to the reference position for the right window.

ELEV	AZIM										
6	-	-	-6	-4	-2	0	2	4	6	8	10
4	-	-8	-6	-4	-2	0	2	4	6	8	-
2	-	-8	-6	-4	-2	0	2	4	6	8	-
0	-10	-8	-6	-4	-2	0	2	4	6	-	-
-2	-10	-8	-6	-4	-2	0	2	4	6	-	-
-4	-10	-8	-6	-4	-2	0	2	4	6	-	-

Table 4-3 Source shift values in degrees with respect to the reference position for the windshield.

ELEV	AZIM								
6	-8	-6	-4	-2	0	2	4	6	8
4	-8	-6	-4	-2	0	2	4	6	8
2	-8	-6	-4	-2	0	2	4	6	8
0	-8	-6	-4	-2	0	2	4	6	8
-2	-8	-6	-4	-2	0	2	4	6	8
-4	-8	-6	-4	-2	0	2	4	6	8

4.3.1 Experimental investigation

A listening experiment was carried out in order to determine MAAs. Stimuli were presented to subjects over two Focal Alpha 80 loudspeakers placed as shown in the Fig. 4-7. Distance between loudspeaker and centre of the listener head was 1.2m. The experiment took place in a semi-anechoic chamber, where the noise floor was below 15 dBA The crosstalk was cancelled according to the method described in [54].

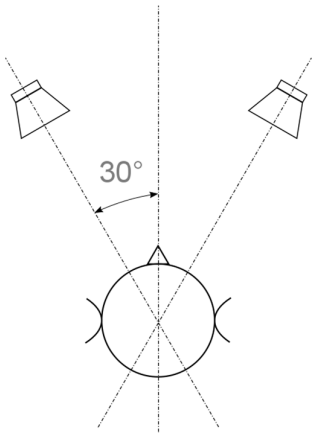


Fig. 4-7 Loudspeakers position for the listening experiment.

Since the objective of the experiment was to obtain MAAs for 3 wind noise sources when they act separately and when they act together, the experiment was divided in 7 short sessions. Table 4-4 contains detailed description of each session. The first three sessions were designed to determine MAA for separated source, as it was discussed in Section 4.2. Sessions no. four to seven, were designed to investigate MAA in the presence of other sounds, in fixed 0° elevation and azimuth position.

Ten listeners without known hearing problems (all males, ages 22–31) participated in the experiment. In the experiment, sounds were presented in pairs. Each pair consisted of a reference sound and a testing sound. The reference sound was a sound appearing in forward direction (0° elevation and 0° azimuth angle). Testing sounds varied in median and horizontal plane according to values presented in Table 4-1, Table 4-2 and Table 4-3. Testing sounds were presented in decreasing angle order, e.g. from 6° to 4°. If the listener perceived a difference three consecutive times, another test sound which is closer to the reference was presented. MAA is defined as the minimum angle at which the listener can perceive a difference between the testing sound and the reference. The order of appearance of the reference sound and the sound under investigation was randomized. The reference signal including the left mirror noise was calibrated at 45dBA (for 0° elevation and azimuth). Level differences between other stimuli are due to the source actual position and to the number of sources acting at the same time.

Table 4-4 Sessions description for MAA free-field experiment. White fields points sources under investigations, grey fields points "maskers" or no source present.

Session number	Left mirror	Wipers	Right mirror
1	For azimuth at 0° elevation For elevation at 0°, -8° and 6° azimuth	-	-
2	-	For azimuth at 0° elevation For elevation at 0° azimuth	-
3	-	-	For azimuth at 0° elevation For elevation at 0°, -8° and 6° azimuth
4	For azimuth at 0° elevation For elevation at 0° azimuth	-	Fixed source at 0° elevation and azimuth
5	For azimuth at 0° elevation For elevation at 0° azimuth	Fixed source at 0° elevation and azimuth	Fixed source at 0° elevation and azimuth
6	Fixed source at 0° elevation and azimuth	Fixed source at 0° elevation and azimuth	For azimuth at 0° elevation For elevation at 0° azimuth
7	Fixed source at 0° elevation and azimuth	For azimuth at 0° elevation For elevation at 0° azimuth	Fixed source at 0° elevation and azimuth

4.3.2 Results and discussion

Fig. 4-8, Fig. 4-9, Fig. 4-13 and Fig. 4-14 present results achieved in the listening experiment for every measurement position averaged over all participants. Polar plots radius represents scores of the experiment within values from 0 to 1. Zero means nobody perceived a difference between the testing sound and the reference, one means all participant perceived a difference. A value of 0.5 was taken (50% of participants perceived a difference) as a threshold for MAA estimation. Bolded lines indicate reference stimuli positions. On polar plots, results are presented in order of appearance in horizontal plane in car cabin i.e. left mirror, wipers and right mirror respectively.

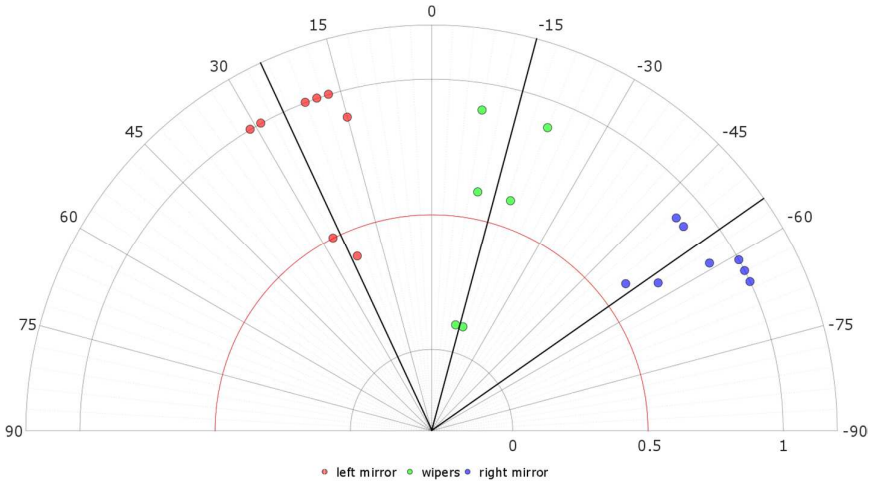


Fig. 4-8 Listening experiment results for azimuth for left mirror, wiper and right mirror at 0° elevation.

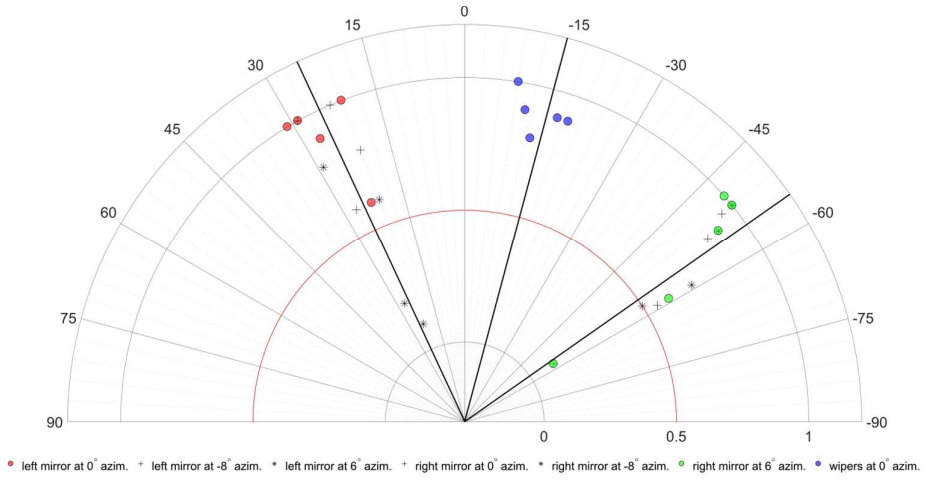


Fig. 4-9 Listening experiment results for elevation for left mirror, wiper and right mirror at different azimuth angles.

In scenarios with only one source present at a time (Fig. 4-8 and Fig. 4-9), MAAs are essentially similar to results achieved using simulated IRs (Fig. 4-3). Of course, previous results cannot be compared directly due to differences in reference position, but values of MAA are still very small. For the left mirror and for the median plane, MAA was found to

be equal 2° for upper positions, and 4 ° for lower positions of the source (Fig. 4-10). For the vertical plane at 6° and -8° azimuth as well as for the horizontal plane, MAA is smaller than measurement accuracy i.e. 2°. Values for the right mirror for the horizontal plane are 4° and 2° for left and right source positions. For the vertical plane only at 6° azimuth MAA is greater than the measurement accuracy (Fig. 4-11). Results for the wiper noise source positions are similar to the right mirror results with only small difference at the horizontal plane (Fig. 4-12).

Valuable information for sound localization techniques could be MAA represented in distance unit. Due to the fact that localization error is equally probable for horizontal and vertical plane, the worst case, i.e. the minimum MAA, should be considered. For only one source present at a time, requirement for sound localization techniques are very demanding. MAA equal to 2° corresponds to approximately 2cm of a source displacement at the left mirror and a little bit more than 3cm for the right mirror. The minimal range of MAA is for instance in the horizontal plane for the left mirror and it is less than 4cm. For the right mirror requirements are a bit less challenging: 6cm for the elevation.

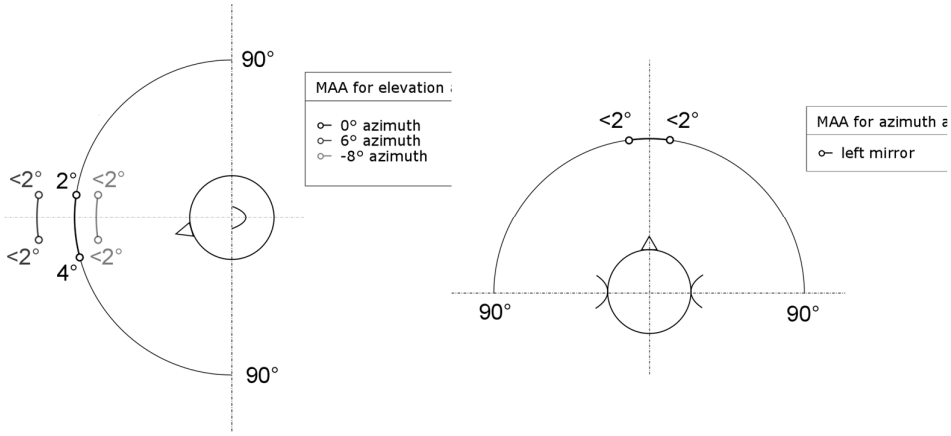


Fig. 4-10 MAA for the left mirror for the horizontal direction (right) and for the vertical direction (left). For MAA on the median plane (left) results are marked with black line and with a grey line for the vertical plane for 6° azimuth and -8° azimuth.

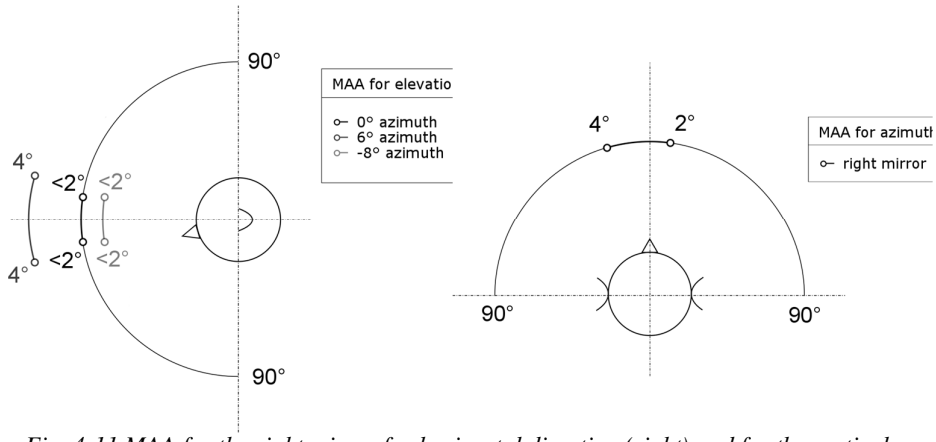


Fig. 4-11 MAA for the right mirror for horizontal direction (right) and for the vertical direction (left). For MAA on the median plane (left) results are marked with black line and with a grey line for the vertical plane for 6° azimuth and -8° azimuth.

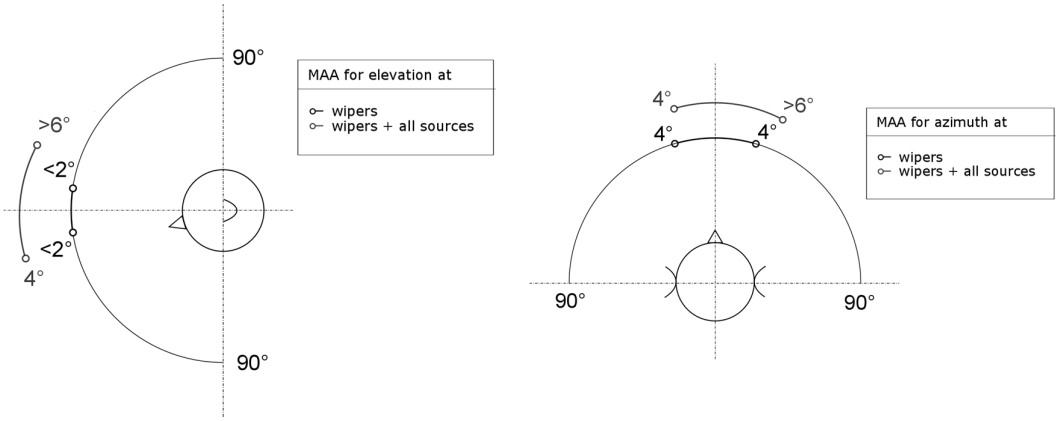


Fig. 4-12 MAA for the wipers for the horizontal direction (right) and for the vertical direction (left). For MAA on the median plane (left) results are marked with black line. With a grey line are marked results with left and right mirrors acting at the same time.

According to predictions, introducing other sound sources while determining MAA changes final results. As can be seen on Fig. 4-13 and Fig. 4-14, results are significantly different compared to the previous scenario. Fig. 4-15 and Fig. 4-16 present the estimated MAA for scenarios with and without further noise sources added. For the left mirror for the median plane MAA stays unchanged (Fig. 4-15). For the vertical plane adding other noise sources makes the value of MAA to increase from 2° to 6° for left direction (towards car cabin

interior). Changes towards windshield are perceived just as without noise. This behaviour can be explained by strong reflections appearing in the windshield corner. Wipers noise MAA as presented in Fig. 4-12 are slightly changed in horizontal plane, but significantly in the median plane. Range of MAA increased from 4° to 10°. The biggest influence of the presence of other noise sources can be observed for right mirror results. The range of MAA increased from 4° to 10° for the median plane and from 6° to 16° for the horizontal plane. This high difference is a consequence of different distance between the driver and left right mirrors. Left mirror generates louder noise at the drivers' position, therefore right mirror noise can be partially masked.

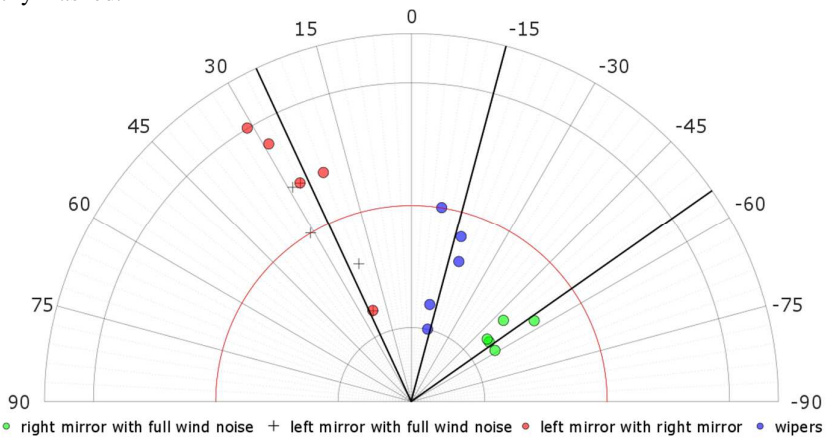


Fig. 4-13 Listening experiment results for azimuth for left mirror, wiper and right mirror with other sources acting at the same time.

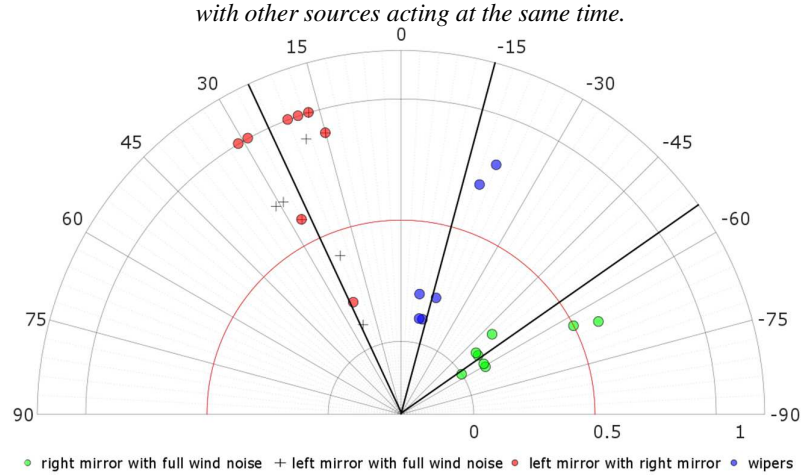


Fig. 4-14 Listening experiment results for elevation for left mirror, wiper and right mirror with other sources acting at the same time.

Masking effect is good news for sound source localization accuracy problem. For only one source present at a time, the minimal range of MAA is, for instance in the horizontal plane, less than 4cm (left mirror). With wiper and right mirror noise added range of unnoticed source displacement raised to 8 cm. For the right mirror requirements are bit less challenging: 6cm for the elevation. Masking has the highest impact on the right mirror noise localization accuracy, where maximal possible error increased from 6cm (median plane) to more than 15 cm.

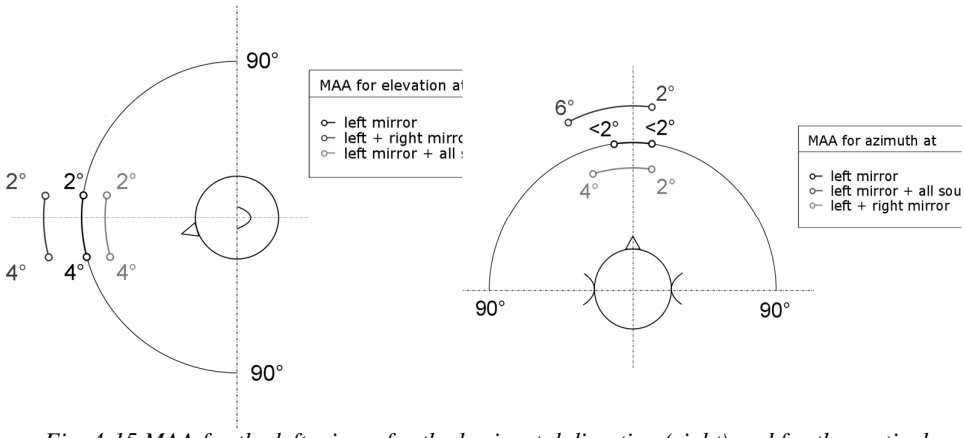


Fig. 4-15 MAA for the left mirror for the horizontal direction (right) and for the vertical direction (left). For MAA on the median plane (left) results are marked with black line. With a grey line are marked results with other sources acting at the same time.

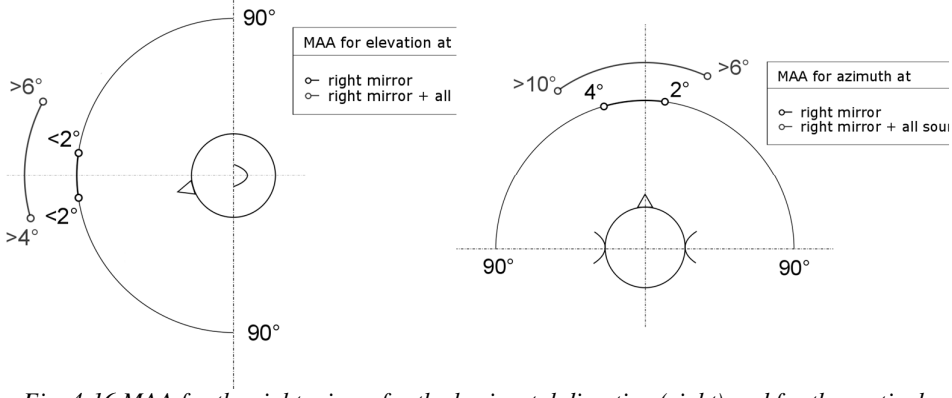


Fig. 4-16 MAA for the right mirror for the horizontal direction (right) and for the vertical direction (left). For MAA on the median plane (left) results are marked with black line. With a grey line are marked results with other sources acting at the same time.

4.4 Conclusions

In this chapter a human-oriented approach to sound sources localization accuracy in car cabins was investigated. A data set of simulated sources evenly distributed on the windshield with 2° angle resolution within the range of 40° from the left to the right with respect to the receiver's position were used in the experiment. The receiver was placed at the centre of driver's head. Minimum Audible Angles were derived from a listening experiments for two vertical planes and the horizontal plane. MAA for the horizontal plane varies from 2° towards left to 4° towards right with respect to the head orientation. For the vertical plane at 0° azimuth (median plane) MAA is approximately 2° higher than in the horizontal plane. MAA for vertical plane is 2° better when a source is moved slightly towards MAA for the horizontal plane. The investigation performed suggests that listeners certainly can perceive a difference when a source is localized farther than 3 cm with respect to its true location. In some cases, changes as small as 1.5 cm can be perceived as well. These values can be interpreted as the accuracy required for a source localization technique if its results are aimed to be exploited for auralization purposes.

Chapter 5

In-vehicle auralization based on statistical methods

5.1 Overview on RIR modelling

Current studies [8] have shown that even in the case of very accurate boundary and source data for simulations combining finite element methods and geometrical acoustic methods, a precise prediction of the sound field in car cabins remains difficult for frequencies far above the Schroeder frequency (in car cabins ≈ 400 Hz). Due to diversity of materials, possible driver's positions and small variations in the car cabin geometry lead to a completely stochastic reformation of the cavity eigenmodes. The goal of an acoustic simulation for frequencies far above the Schroeder frequency can therefore only be the best possible prediction of the sound decay and sound energy in all frequency bands [8]. This chapter aims at presenting a novel approach, based on a statistical model, to perform auralization in car interior applications. The interest in this research field is driven by the necessity of reducing the use of wind tunnel tests for assessing the impact of wind noise in vehicle interior acoustics.

Statistical modelling of measured RIRs is used in many research topics such as dereverberation [24, 55, 56, 57] or reverberation parameter estimation [25, 58]. After Sabine's first calculation method of the reverberation time, 50 years later Schroeder extended Sabine's fundamental work [13] and derived a set of statistical properties describing the frequency response of a random impulse response. He proposed a method in which the room impulse response, for a given source and receiver position, is given by a single realization of a non-stationary stochastic process. The model gives a two-parameter representation of the impulse response, and is valid for late reverberations and frequencies greater than the Schroeder frequency [59]. In [24] a model, which is an extension of a Polack's model, valid for some early reflections was proposed. The approach consisted in dividing the RIR into two segments with different parameter values. Recently, yet another extended model of RIR, which represents both early and late reverberation components in multiple frequency bands, was presented [60]. All these approaches are well suited for room acoustics. When moving

to car interior acoustics, on the contrary, these models lack of sufficient accuracy, since car cabins are characterized by a peculiar acoustic environment, in which reverberation time is short and reflections appear close to each other in time. Also as it was presented in Chapter 3, perception of an acoustic environment inside a car strongly depends on early reflections. Therefore, a new statistical model focused on in-vehicle applications is proposed with the main objective of deriving an auralization procedure that creates wind noise signals perceptually indistinguishable from real measurements without providing any geometry of the car under investigation. In order to tune the model for in-vehicle purposes, measured car RIRs are used to estimate correct values of model parameters like RT60, DRR, ETL and EDT. RIRs are measured on evenly distributed points surrounding the main wind noise sources i.e. left and right mirrors, A-pillars and wipers. The procedure is repeated for a sufficient number of cars in order to get a dataset of RIRs and corresponding values of model parameters. This allows to adapt the model according to achieved results and to propose a small number of RIR measurement points for the end user of the model. Thanks to this approach a time-saving objective of the proposed method will be fulfilled. Instead of measuring RIRs for every desired source position, only a few measurement points will be necessary. Rest of the required data will be generated by the model allowing virtually relocate sources for auralization. It will be an exceptional advantage in performing so-called what-if analysis.

5.2 New statistical model

In Chapter 1 the statistical model developed by Polack [5] was introduced. It gives a good approximation of RIRs for times greater than the mixing time. Therefore, the main drawback of this model is the fact that it can be used only to simulate late reverberation. Also existing extensions of the Polack's model are not suitable for simulating RIRs of a car cabin. The generalized model of [30] introduces a second Gaussian decaying noise for the early part of RIR. The other model reported in [31] combines DRR and a drop in energy after the direct sound in Polack's model. Both of them are focused on rather bigger enclosures than car cabins. The presence of very strong early reflections requires a better treatment of the early part of RIRs. Therefore, a new statistical model is proposed here, incorporating the Direct to Reverberant energy Ratio (DRR), Early to Late Index (ETL), Early Decay Time (EDT) and Reverberation Time (RT60).

A graphical overview of the proposed method is presented in Fig. 5-1. In order to auralize in-vehicle wind noise four steps are required. In the first step sound sources are precisely localized using one of the state-of-the-art SSL techniques. An interesting approach for SSL which can be applied with proposed model is the time domain inverse beamforming [61]. This method is particularly interesting because can give a time signal of a localized source.

Other SSL method that can be utilized for the purpose of the model is for instance acoustic holography. Preferably, selected method should meet requirements depicted precisely in Chapter 4. With source location information direct sound of a source can be auralized in a simple way by taking a Dirac pulse and convolve it with HRTF filters corresponding to the source position.

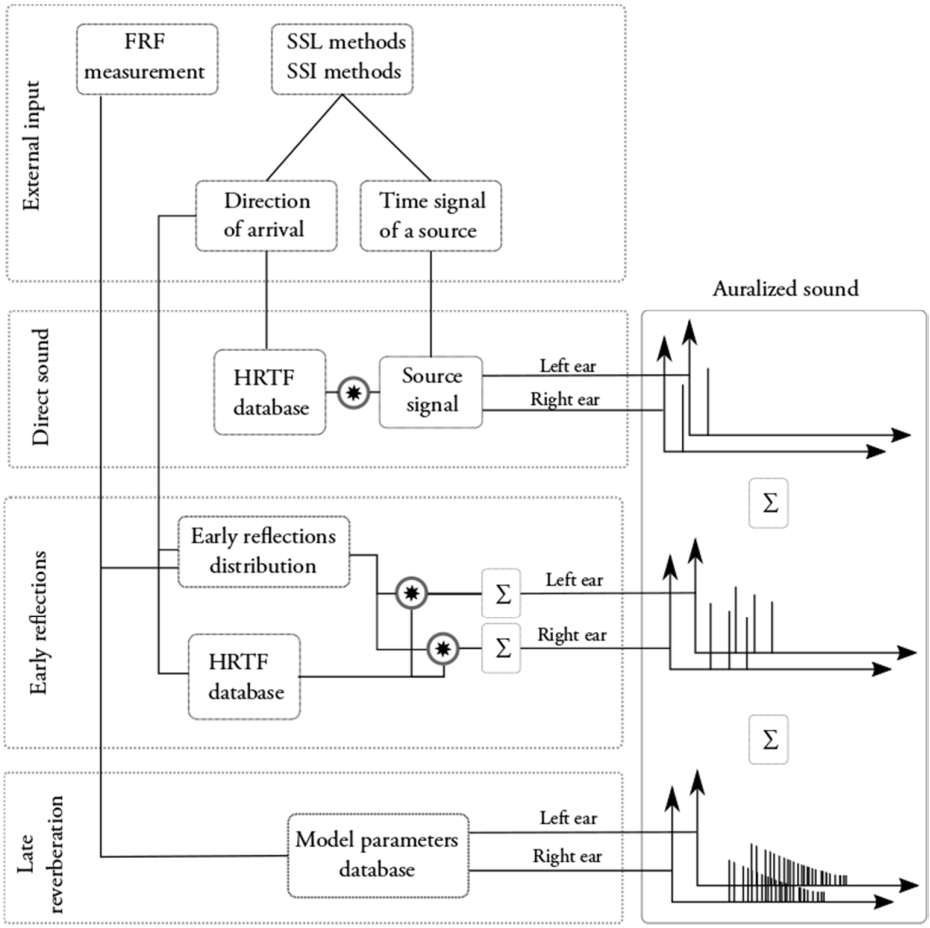


Fig. 5-1 Block diagram of the processing required for the auralization based on the proposed method

As mentioned before, early part of the RIR needs to be modelled with great care. Omitting strong early reflections can cause different to colouration of an auralized sound and can be perceived by listeners, as shown in Chapter 3. One of the objectives of this method is to derive a good approximation of RIRs without using any car model and therefore without exploiting any BEM, FEM or GA method. For a dynamically changing environment it seems to be unrealistic task. Different source and receiver positions, changing in time, require constant update of RIRs. However, the purpose of the new model is wind noise auralization, in which the main sources are localized in rather known positions. As presented in Fig. 1-4, the main wind noise sources are located around door side mirrors, wipers and A-pillars. This source distribution in a car cabin does not change drastically among cars. Therefore, it can be assumed that, inasmuch as the sources are in fixed positions, early reflections will exhibit the same behavior. RT models of different cars were used in order to estimate the most probable area of early reflection appearance for each wind noise source. As a result, auralization model uses early reflection directions and time of arrivals for a "statistical car". Later it will be shown how amplitudes of early reflection are calculated.
Late reverberations are estimated in similar as it was proposed by Polack. However in the new method more parameters are used including DRR, ETD and ETL on top of the Polack's mode. As a consequence, the energy of late reverberations depends on the energy of a direct found and early reflections, which varies with source position. Therefore a scaling factor is introduced for late reverberations. Below the mathematical description of the new model will be given.

The original Polacks' model is described as:

$$h(t) = \begin{cases} b(t)e^{-\zeta t}, & t \geq 0; \\ 0, & t < 0, \end{cases} \qquad 5.1$$

where *b(t)* is a zero-mean stationary Gaussian noise, and ζ is a damping constant related to the reverberation time RT60 by the equation 1. 2.

The new statistical model can be expressed as

$$h_{r,k}(t)=\begin{cases} HRIR_r * \delta(t), & t = t_d; \\ \sum\left[HRIR_{j,r} * \delta(t - t_{j,r})\right]\ e^{-\xi_e t}, & t_{mix} \geq t > t_d; \\ b_{cor}(t)e^{-\xi_l t}, & t \geq t_{mix}; \\ 0, & t < t_d, \end{cases} \qquad 5.2$$

where

$HRIR_r$ - Head Related Impulse Response corresponding to the direct sound,

$\delta(t)$ - Dirac delta function for the direct sound,

$HRIR_{j,r}$ - Head Related Impulse Response corresponding to the *j*-th reflection,

$\delta(t - t_{j,r})$ - Dirac delta function for the *j*-th reflection,

ζ_e, ζ_l - damping constants for the early reflections and the late reverberation, respectively, defined as:

$$\zeta_e = \frac{\log(10^3)}{EDT}, \qquad 5.3$$

$$\zeta_l = \frac{\log(10^3)}{RT_{60}}, \qquad 5.4$$

$b(t)_{cor}$ - random Gaussian noise including a correction factor based on DRR and ETL expressed by

$$b_{cor}(t) = \frac{\int_{t_{mix}}^{\infty} h_{cor}{}^2(t)}{\int_{t_{mix}}^{\infty} h^2(t)} b(t). \qquad 5.5$$

The integral in the numerator of the equation above needs to be clarified. In order to correct energy of direct sound, early and late reverberation part parameters DRR and ETL should be included. The integral of $h(t)_{cor}$ can be interpreted as a required energy of the late reverberation part of $h(t)$ after the correction. Below the integral will be derived.

In Chapter 1 two parameters characterizing RIRs were introduced, DRR and ETL given by equations :

$$\mathrm{DRR} = 10\log_{10}\frac{\int_{0}^{t_d} h^2(t)}{\int_{t_d}^{\infty} h^2(t)}, \qquad 5.\ 6$$

$$\mathrm{ETL} = 10\log_{10}\frac{\int_{0}^{t_{mix}} h^2(t)}{\int_{t_{mix}}^{\infty} h^2(t)}. \qquad 5.\ 7$$

The total energy of the RIR can be expressed by sum of two integrals corresponding to the direct path and the reverberant path

$$\int_{0}^{\infty} {h_{cor}}^2(t) = \int_{0}^{t_d} h^2(t) + \int_{t_d}^{\infty} h^2(t). \qquad 5.\ 8$$

In equation 5. 2 we introduced division of the reverberant component in two separate components describing early reflections and late reflections. Therefore, equation 5. 3 can be rewritten as

$$\int_{0}^{\infty} {h_{cor}}^2(t) = \int_{0}^{t_d} h^2(t) + \int_{t_d}^{t_{mix}} h^2(t) + \int_{t_{mix}}^{\infty} h^2(t) \qquad 5.\ 9$$

This form of the equation 5. 8 implies the new form of the ETL equation

$$\mathrm{ETL} = 10\log_{10}\frac{\int_{0}^{t_d} h^2(t) + \int_{t_d}^{t_{mix}} h^2(t)}{\int_{t_{mix}}^{\infty} h^2(t)}. \qquad 5.\ 10$$

Rearranging the equation above gives

$$10^{0.1\mathrm{ETL}} \int_{t_{mix}}^{\infty} h^2(t) = \int_{0}^{t_d} h^2(t) + \int_{t_d}^{t_{mix}} h^2(t) \qquad 5.\ 11$$

$$\int_{t_d}^{t_{mix}} h^2(t) = 10^{0.1\mathrm{ETL}} \int_{t_{mix}}^{\infty} h^2(t) - \int_{0}^{t_d} h^2(t) \qquad 5.\ 12$$

Similar operation can be done with the equation

$$\int_{0}^{t_d} h^2(t) = 10^{0.1\mathrm{DRR}} \int_{t_d}^{\infty} h^2(t)\,. \qquad 5.\ 13$$

By combining equation 5. 11 with the equation 5. 12 we obtain the total energy od the RIR including ETL parameter

$$\int_{0}^{\infty} {h_{cor}}^2(t) = \int_{0}^{t_d} h^2(t) + 10^{0.1\,\mathrm{ETL}} \int_{t_{mix}}^{\infty} h^2(t) - \int_{0}^{t_d} h^2(t) + \int_{t_{mix}}^{\infty} h^2(t)\,, \qquad 5.\ 14$$

$$\int_{0}^{\infty} {h_{cor}}^2(t) = (1 + 10^{0.1\,\mathrm{ETL}}) \int_{t_{mix}}^{\infty} h^2(t)\,. \qquad 5.\ 15$$

Using the relation between total RIR energy and the direct path energy equation 5. 13 can be written as

$$10^{0.1\mathrm{DRR}} \left(\int_{0}^{\infty} {h_{cor}}^2(t) - \int_{0}^{t_d} h^2(t) \right) = \int_{0}^{t_d} h^2(t)\,, \qquad 5.\ 16$$

and simplified to the form

$$\int_{0}^{\infty} {h_{cor}}^2(t) = \left[1 + \frac{1}{10^{0.1\,\mathrm{DRR}}} \right] \int_{0}^{t_d} h^2(t)\,. \qquad 5.\ 17$$

Because equations describe sum of the same RIR, we can merge right sides of the equations and write

$$(1+10^{0.1\,\mathrm{ETL}})\int_{t_{mix}}^{\infty} {h_{cor}}^2(t) = \left[1+\frac{1}{10^{0.1\,\mathrm{DRR}}}\right]\int_0^{t_d} h^2(t)\,. \qquad 5.\ 18$$

The total energy of the late reverberation part in therefore

$$\int_{t_{mix}}^{\infty} {h_{cor}}^2(t) = \frac{1}{1+10^{0.1\,\mathrm{ETL}}}\left(1+\frac{1}{10^{0.1\,\mathrm{DRR}}}\right)\int_0^{t_d} h^2(t)\,. \qquad 5.\ 19$$

The integral from equation above can be use directly for equation 5. 5 resulting a scaled late reverberation based on parameters DRR and ETL.

5.3 Parameters estimation

The model uses a statistical distribution of key RIR parameters and early reflections. The model specified by equation 5. 2 requires four input parameters: RT, DRR, EDT, ETL. Two different type of RIRs are provided: based on measurement in real cars and ray tracing simulation using car models. Both methods have advantages and disadvantages, therefore combining them together in RIR database, gives a better overview on car cabin acoustics.
Calculations of RIR's parameters were performed using Matlab script. RT and EDT were calculated using a method based on ISO 3382-1:2009 using reverse cumulative trapezoidal integration to estimate the decay curve, and a linear least-square fit to estimate the slope between 0 dB and -60 dB or -10 dB for EDT. DRR and ETL were calculated based on equations 1. 6 and 1. 7. Statistical analysis of the data were carried out using Statistica and R environment.

5.3.1 Measured RIRs

Obtaining good measurement RIRs from a source located on a surface and, moreover, when source-receiver distance is relatively small, is not an easy task. The sound sources have to be omnidirectional and have a negligible size in order not to influence the field, especially in the higher frequency range. Additionally, required spatial resolution of acquired RIRs demands a very small source dimension. Calibrated LMS small volume acceleration sources are perfectly suited for this purpose. The shape of the source allows attaching it directly to a flat surface. The diameter of the source is approximately 3cm, which meets requirements for

the proposed measurement campaign. The measurement set-up is presented in Fig. 5-2 and Fig. 5-3.

In order to have a full overview of RIR in areas of possible wind noise locations, sources were located on windshield, left door window and right door window. Measurement were designed to investigate even small influences of source position on RIR, therefore sources were distributed evenly by creating a 10x10cm grid. The G.R.A.S. KEMAR head and torso simulator placed at the driver's position was used to acquire binaural RIR. To control possible HATS displacement, which could affect results, HATS was controlled by two laser beams pointed at the left ear.

Fig. 5-2 Measurement set-up for BRIR acquisition from sources located on a windshield.

Using the small volume acceleration source has one more additional advantage. The source has an internal sound source strength sensor which outputs a real-time volume acceleration signal. By acquiring the output of this sensor a very accurate RTFs can be calculated. The acquisition was performed using an LMS SCADAS frontend together with LMS Test.Lab software. RIRs were calculated based on measured RTFs by applying IFFT on measured RTF and its complex conjugate. The excitation signal was a 50% burst random noise with frequency range 1-10kHz.

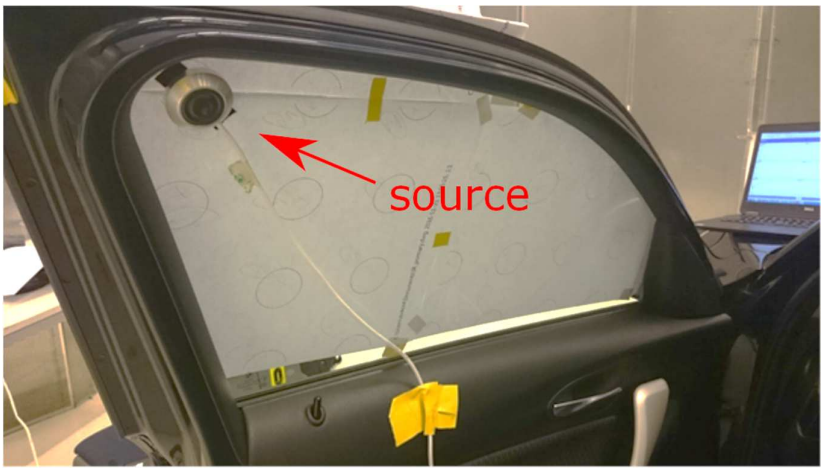


Fig. 5-3 Miniature volume acceleration source position of a left door window

Measurements of RTF in real cars were performed in a semi anechoic chamber. Four different cars were tested: a compact size hatchback, two compact size estate cars and one minivan.
The measurement set-up presented above, even though it is very well suited for obtaining wind noise RIRs, has some limitations as well. The miniature source, due to the small dimensions, is limited frequency range. The nominal lower frequency of this source is 2kHz. With lower excitation signal amplitude, the frequency range can be extended to 1kHz at the cost of a lower SNR. Fortunately, 1kHz is a sufficient frequency for wind noise reproduction. SNR, as will be presented later, was also high enough to correctly estimate RIR parameters.

5.3.2 Simulated RIRs

The second dataset of RIR comes from RT calculations. Seven models of different car sizes and types were investigated: two mid-class sedans, one mid-class estate, one small city car, two SUV cars and one sport coupe. In all models, individual cavity parts like dashboard, windows, seats have the same acoustic impedances. Thus it is possible to focus mostly on geometrical differences. Sources and the receiver were placed at the same positions as used for the measurement in real cars, in order to obtain comparable results. The total number of points depends on car geometry, and varies from approximately 120 to 150. Fig. 5-4 illustrates source distribution on car windshield and door windows.
RT calculations were performed using Virtual.Lab Ray Tracking software. Outputs of RT techniques are echograms for each simulation point. They are later convolved with HRTFs

selected according to directions of arrival of each ray path, in order to give spatial information to RIR. The KEMAR HRTF database from MIT is used for convolution [34].

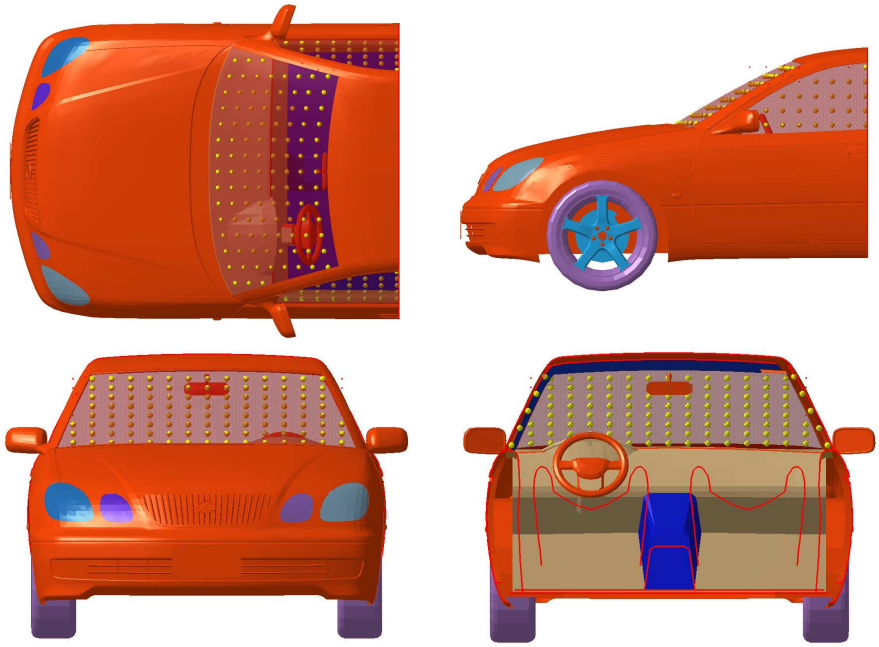


Fig. 5-4 An example of sources distribution in a car (yellow points)

5.3.3 RIR clustering

Cars obviously do not have the same size and shape of a body, including windshield and side windows. This fact has an impact on number of RIRs acquired. Measurement and simulation points, even though created from a uniform 10x10cm grid, have different boundary caused by windows shape. To compare results, i.e. RIR calculated parameters from one area of a car with the same area from different car, it is necessary to normalize geometries in order to have common boundary. Fig. ***5-5*** illustrate this very important aspect.

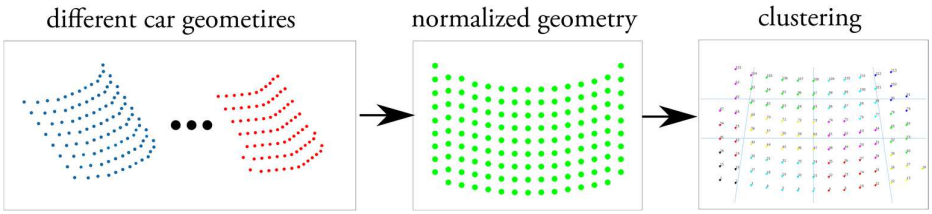


Fig. 5-5 Illustration of necessary steps for geometry clustering

Various clustering algorithms were investigated in order to select the most suitable method. First the K-means clustering method was checked. The algorithm divides measurement points together with assigned parameter values into K clusters so that the within-cluster sum of squares is minimize. Another applied method was the Cross-Entropy Clustering which is a combination of K-means with expectation–maximization method [62]. Both methods group results in clusters with relatively small within-cluster standard deviation of RIR parameters. Unfortunately, they have common disadvantage of creating different number of clusters for different RIR parameters, and sometimes even for the same parameter and different car. Therefore, comparing results from car to car, which is the main purpose of clustering, stays difficult. Eventually, the best suited method is the simplest one. Clusters are created by geometrical division. For the windshield, upper and bottom side was divided in four parts of the same relative length, and left and right sides in three parts. As a consequence, windshields were divided in twelve clusters for each car. An example of clustered windshield is shown in Fig. 5-6.

Clustering of side windows is more straightforward. K-means clustering and Cross-Entropy Clustering method give almost the same results for all cars and all parameters. Each window has two clusters assigned: one in trapezoidal shape located closer to the dashboard, and one in rectangular shape as presented in Fig. 5-7.

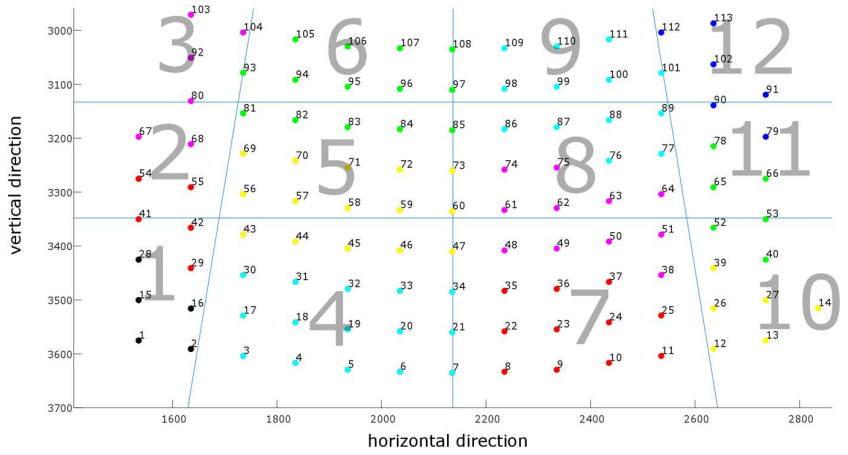


Fig. 5-6 An example of clustering measurement points on a windshield into 12 clusters. Grey numbers indicate cluster numbers and colour of points indicate cluster assignment.

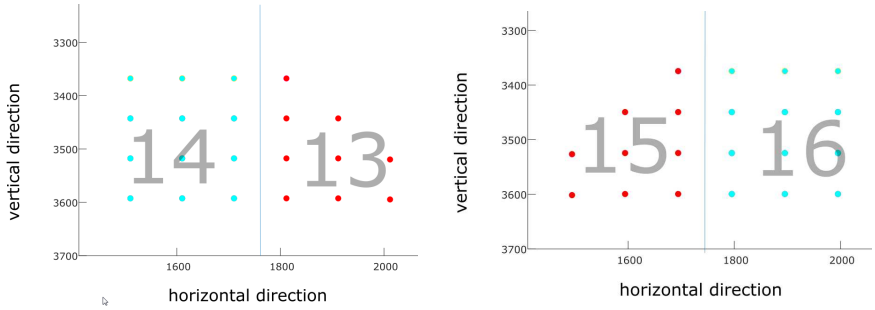


Fig. 5-7 An example of clustering measurement points on a left and right door window into 2 clusters. Grey numbers indicate cluster numbers and colour of points indicate cluster assignment.

5.4 Measurement and simulation results

For each RIR obtained from measurement or simulation, parameters RT60, DRR, EDT, ETL were calculated. In both scenarios binaural receiver was used, which doubles the number of parameters: one set of four parameters for the left and one for the right ear. Due to the fact that RT60 and EDT may depend on frequency, calculations were performed in octave bands. The purpose of obtaining detailed representation of RIRs parameters is twofold. First, mean values for the entire population in each cluster ("statistical car") can be applied directly for

the auralization model. Some parameters may keep the same value for every position in car cabins. The second reason of performed measurements and simulation is to indicate possible statistically significant differences and relations between clusters within the “statistical car”. Statistical test were performed in order to check parameters correlation. It was found that all parameters are uncorrelated with each other. Correlation coefficient is typically below 0,25. It means that parameters carry different information and therefore all of them can be used they can be used and analyzed independently. Fig. 5-8 presents an example of two parameters correlation.

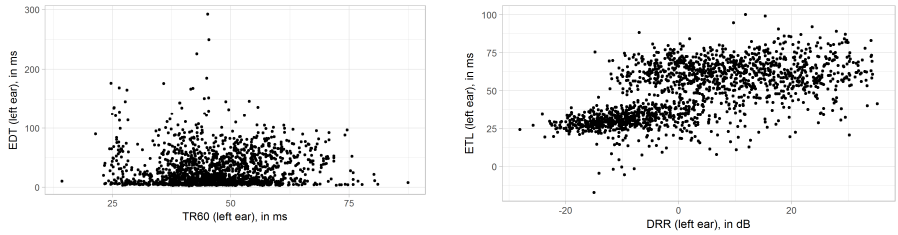


Fig. 5-8 Correlation between RT60 and EDT (left) and DRR and ETL (right) for all cars

Detailed results can be visually inspected by plotting them as shown in Fig. 5-9, Fig. 5-10, Fig. 5-11 and Fig. 5-12 (missing points on the top of the plot are due to rear-view mirror position). These type of results representation give an overview on spatial sensitivity of calculated parameters. Measurement points are then clustered as presented in the section above.

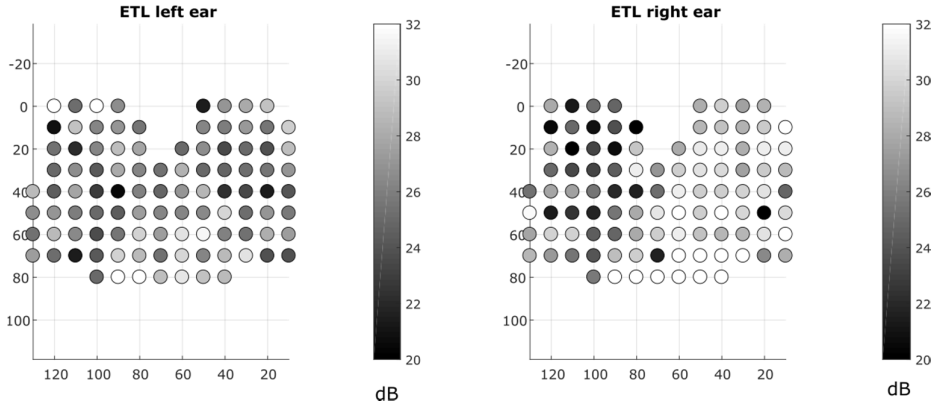


Fig. 5-9 ETL for left and right ear for a single car

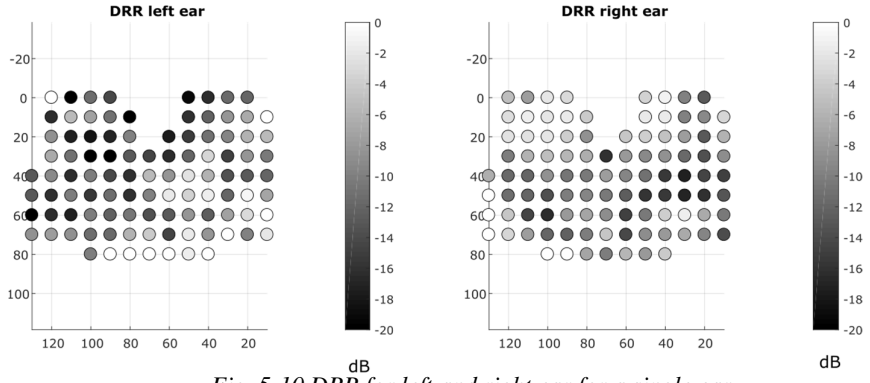
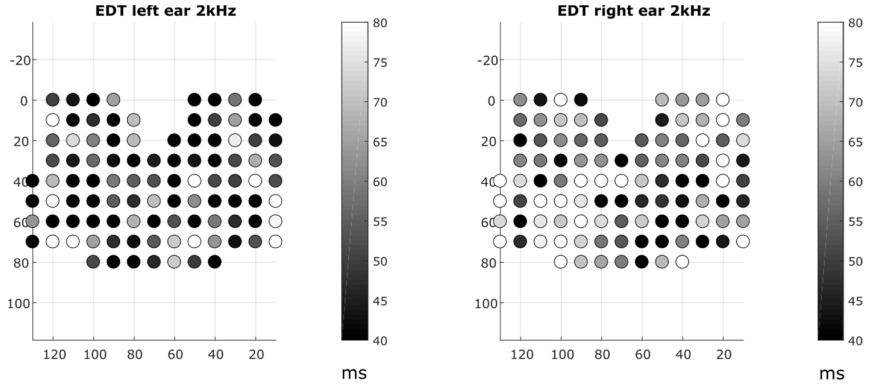


Fig. 5-10 DRR for left and right ear for a single car



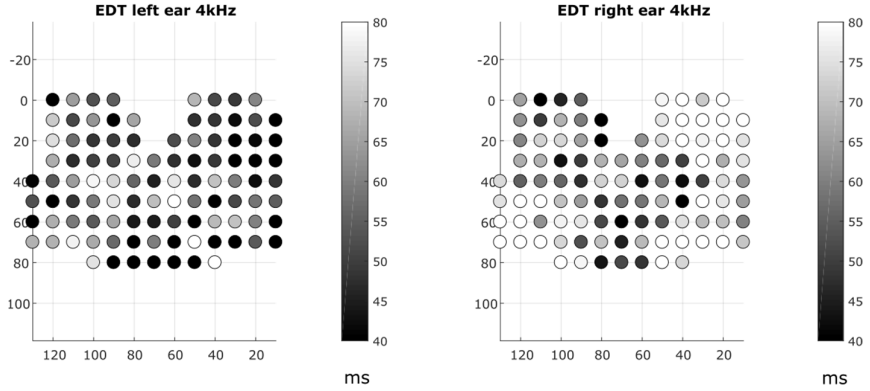


Fig. 5-11 EDT for left and right ear for a single car

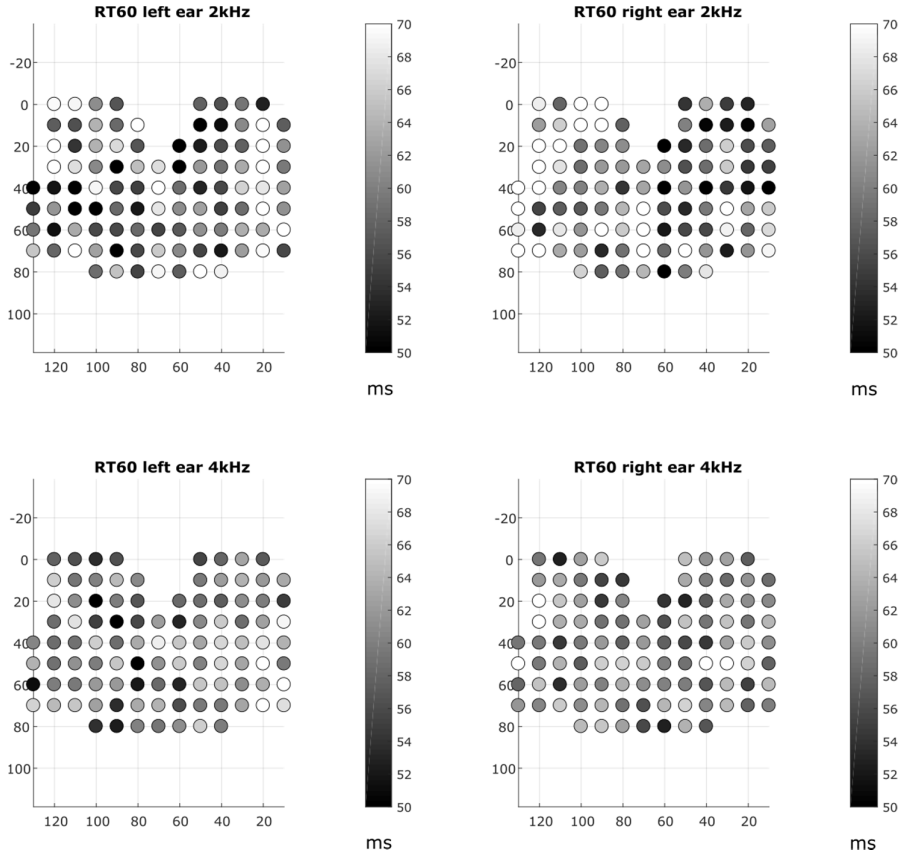


Fig. 5-12 RT60 for left and right ear for a single car

Below, box plots showing values of DRR, ETL, RT60 and EDT for all cars considered in the experiment are presented. For all plots points inside a box indicate the mean value, box size corresponds to the standard error defined as a fraction of the standard deviation to the number of observations, and whiskers shows the standard deviation.

The highest statistical differences between clusters can be observed for DRR and ETL parameters for both ears (Fig. 5-13 and Fig. 5-14). Mostly sources placed on the left door window differs from other clusters (cluster no. 13 and 14). Parameters DRR and ETL give information of energy ratio of early parts of RIR to late reverberation. High values of these parameters indicate significant influence of the direct path and early reflections on RIR.

Therefore is clear that the closest clusters to the receiver have different values of DRR and ETL.

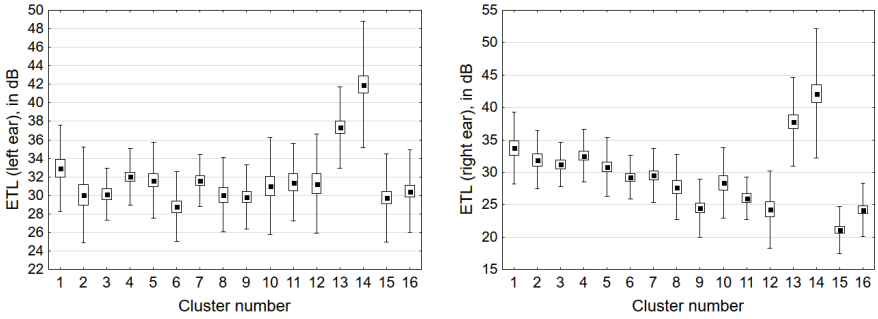


Fig. 5-13 Box plot for ETL for left ear (left) and right ear (right).

Values of ETL for windshield clusters (no. 1 to 12) at a right ear decrease with distance. It is a natural behavior, because with higher source-receiver distance influence of early reflections on RIR is lower. For a left ear ETL varies slightly for clusters close to the reviver and keeps almost the same values for clusters close to a passenger seat.

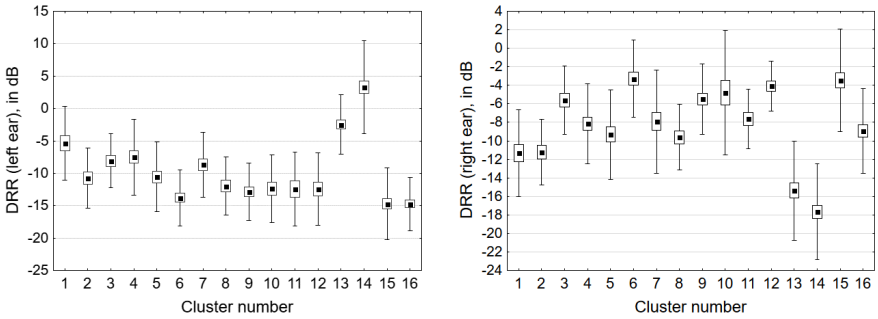


Fig. 5-14 Box plot for DRR for left ear (left) and right ear (right).

Values of ETL, as well as DRR for the left ear in clusters close to a dashboard around driver's position (clusters 1,4,7) have higher values than other cluster on a windshield. Differences are caused by a strong reflection from a dashboard which arrives to the receiver almost at the same time. Due to finite sampling rate of RIR, these reflections events are captured together with a direct path creating one peak. In the same way can be explained high DRR values for clusters close to a rearview mirror for a right ear (clusters 3, 6, 9).

Fig. 5-15 and Fig. 5-16 presents box plots for RT60 and EDT. It can be noticed, that values of those parameters vary less among clusters than for DRR and ETL. Both RT60 and EDT

describe RIR properties in time. Due to relatively small dimensions of cars, difference in source positions have small effect on those parameters. Nevertheless visual inspection of box plots allows to say that for sources located on left window, TR60 and EDT for ipsilateral ear of HATS have significant lower values than for sources located on a windshield and a right window.

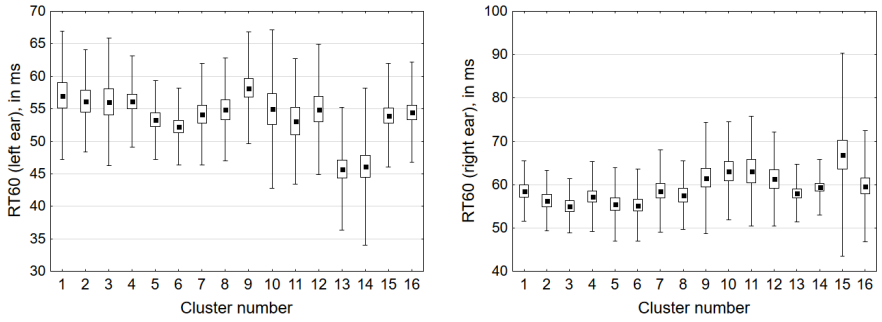


Fig. 5-15 Box plot for RT60 for left ear (left) and right ear (right).

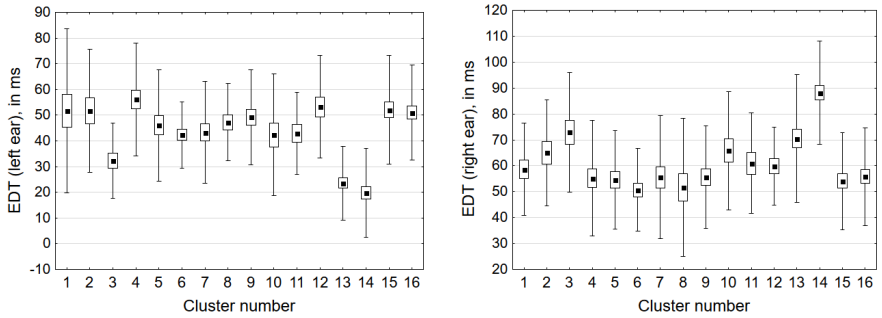


Fig. 5-16 Box plot for EDT for left ear (left) and right ear (right).

For some clusters it is as well possible to predict value of parameters based on correlations with other clusters. An example of cluster correlations is presented in Fig. 5-17 and Fig. 5-18. By measuring RIR in cluster no.3 values for DRR can be predicted in clusters no. 4 and no. 13. This property can reduce number of measurement points necessary to “tune” the statistical model for a correct car.

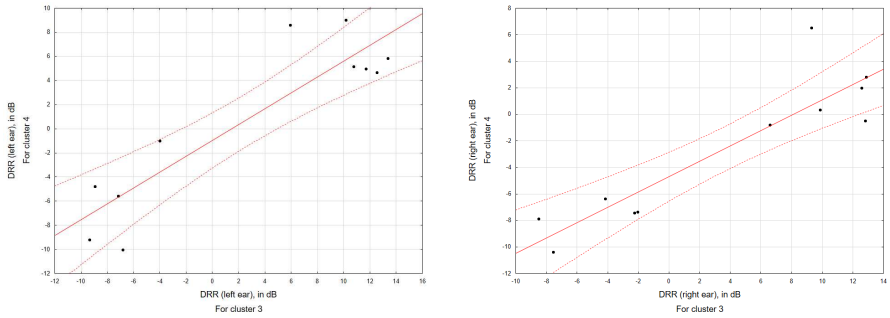


Fig. 5-17 Correlation between cluster no. 3 and cluster no. 4 for DRR (left and right ear)

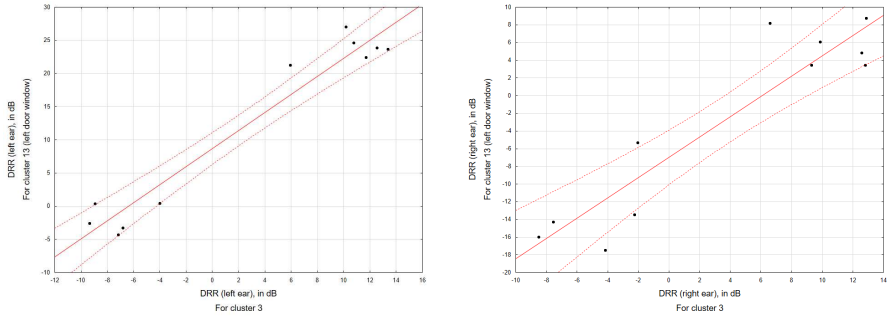


Fig. 5-18 Correlation between cluster no. 3 and cluster no. 13 for DRR (left and right ear)

5.5 Validation of the model

The model was validated by calculating measured and RT simulated RTFs from crucial points corresponding to main wind noise sources (Fig. 5-19) and compering them with FRFs calculated by the statistical model. Seven points were selected for validation: one for left and right mirror noise, one for wiper noise and two for left and right A-pillar noise.

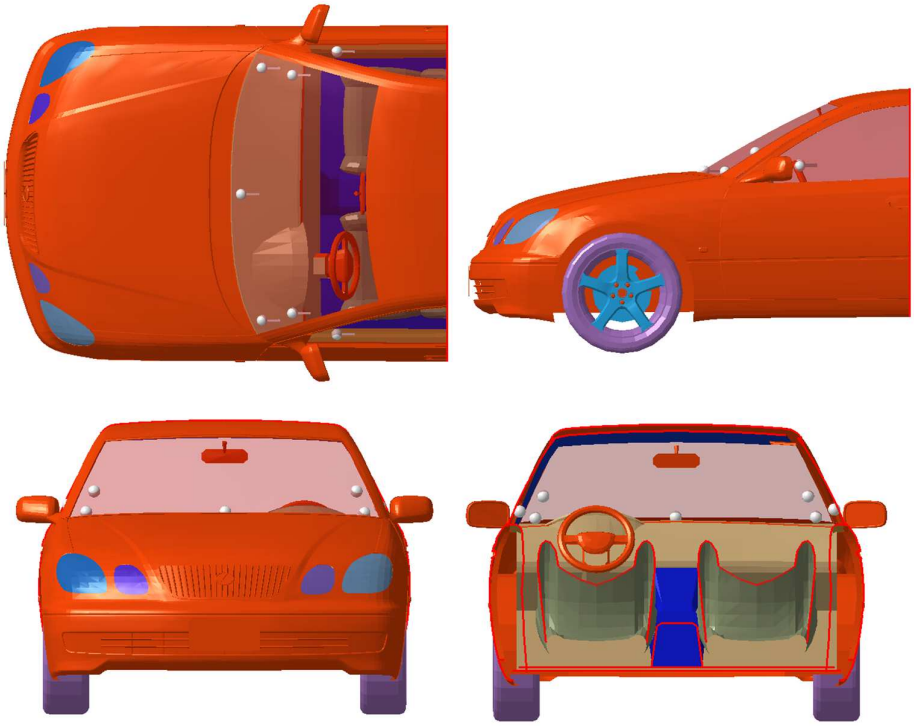


Fig. 5-19 An example of wind noise sources distribution in a car (white points) used for the statistical model validation

Two cars, named in the thesis Car A and Car B were sources of reference RTFs for the statistical model validation. Results are presented in Fig. 5-20 to Fig. 5-47. RTFs are calculated by the statistical model for two input parameters sets: for the "statistical car" and "tuned" parameters. The "statistical car" parameters are averaged parameters over whole measured and simulated population (presented in sec. 5.4) . This method gives only a rough approximation of RTFs. In order to achieve better results, "tuned" parameters should be used. They are taken from at least one measured RTF and based on the knowledge of the statistical

distribution of acoustic parameters in a car (sec. 5.4), missing parameters values are calculated.

For each car and each measurement point, four plots are presented: two for different parameters sets for left and right ear separately.

In all figures presenting RTFs calculated for the "statistical" car one interesting attribute can be noticed: RTFs are more fluctuated with respect to reference signals. It is caused mainly due to DRR and ETL differences. Cars A and B have relatively high DRR and ETL values (20-30 dB) whereas some cars have values of DRR below 0 dB. Using only averaged values for all cars is, as shown in figures below, not sufficient.

For the left mirror noise (Fig. 5-20 to Fig. 5-23) results calculated by the model are very good for both cars. Even using averaged parameters values RTFs look similar to references. With correct values of the model parameters (Fig. 5-21 and Fig. 5-23) RTFs are almost identical as references. In fact, this point is the easiest to model because the direct path plays dominant role. During analysis of RT results for different car models, it was found, that for this source location only one strong reflection might be present (within 10 dB relative to the direct) and occurs later than 1ms after the direct path. Therefore having localized the direct path correctly is essentially sufficient to obtain a correct RTF.

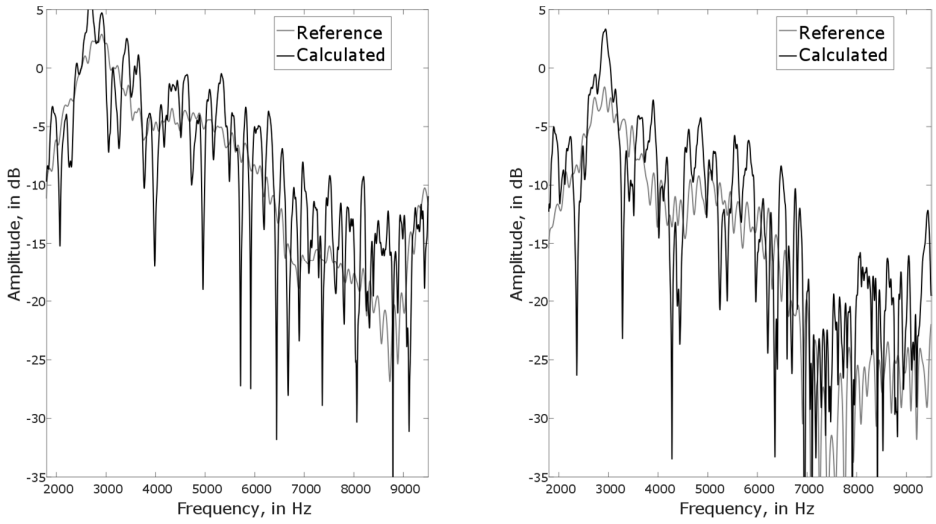


Fig. 5-20 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the left mirror, car A.

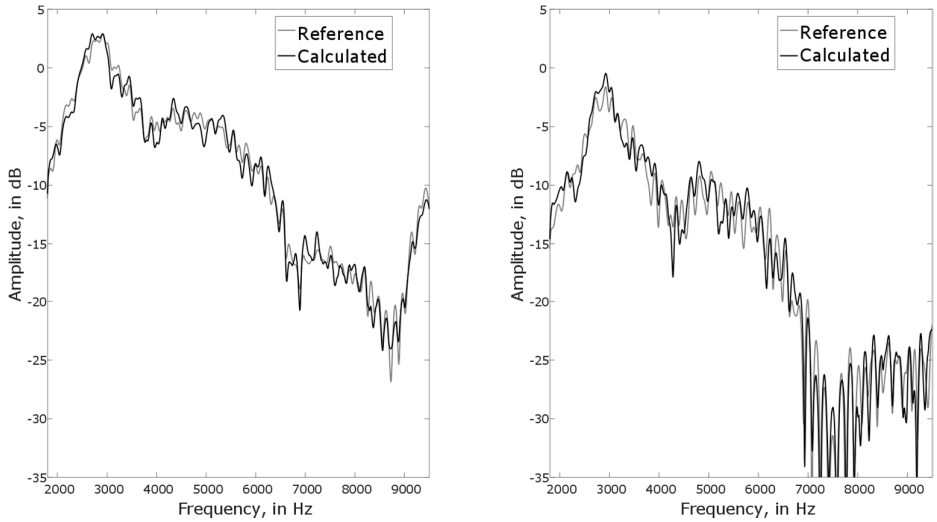


Fig. 5-21 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the left mirror, car A.

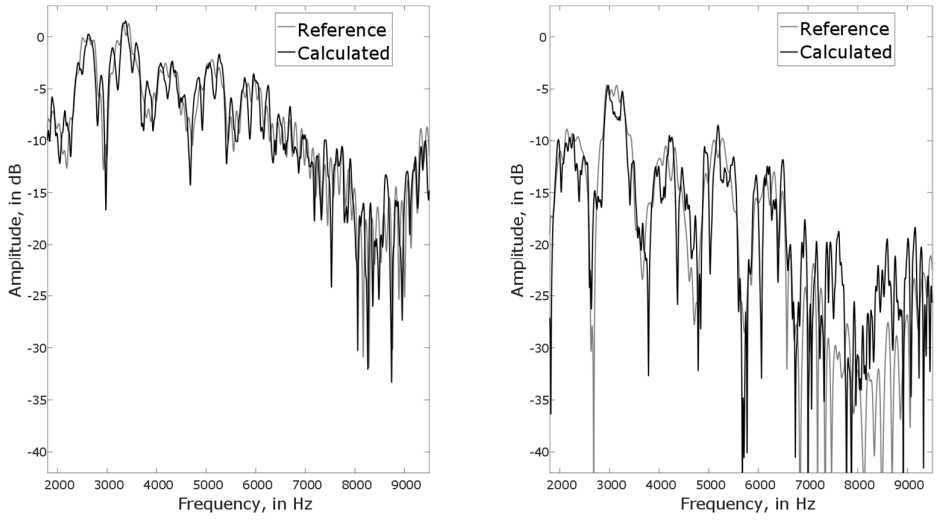


Fig. 5-22 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the left mirror, car B.

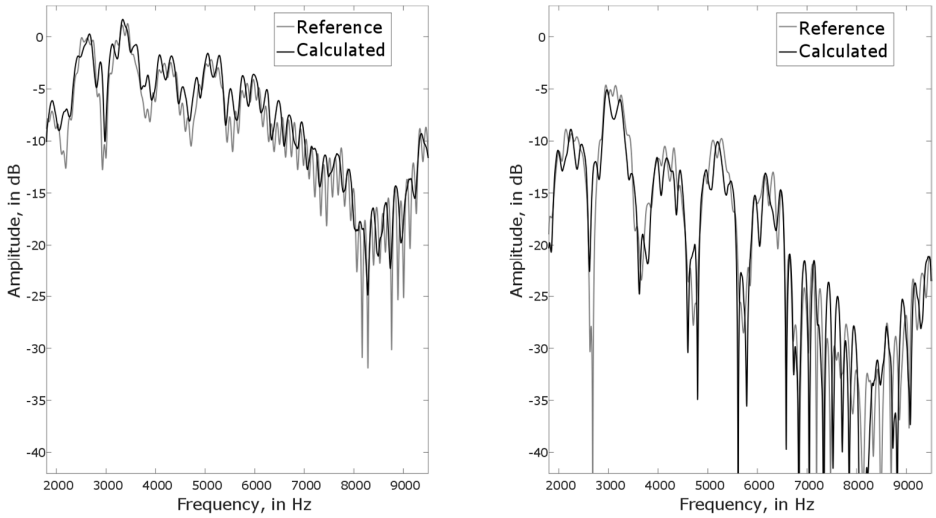


Fig. 5-23 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the left mirror, car B.

Fig. 5-24 to Fig. 5-31 show RTFs calculated by the model for the left A-pillar. For this wind noise source type two points were calculated during validation: the first point very close to the dashboard, and the second point 20cm higher (Fig. 5-19). The purpose of having two validation points for this area is twofold: A-pillar wind noise is rather a distributed source, therefore in practice, two sources are necessary and the source location in the dashboard-windshield corner might be difficult for the statistical model. Indeed, results are not as good as for the left mirror. For the first point for the Car B, RTF only roughly matches the reference as shown in Fig. 5-27.

Differences between calculated and reference RTFs, as mentioned above, are mainly due to strong reflections arriving right after the direct path. Time delays between reflections are sometimes 0.1-0.2ms and vary from car to car. Therefore, it can happen that for some source locations a universal set of early reflections might be not sufficient.

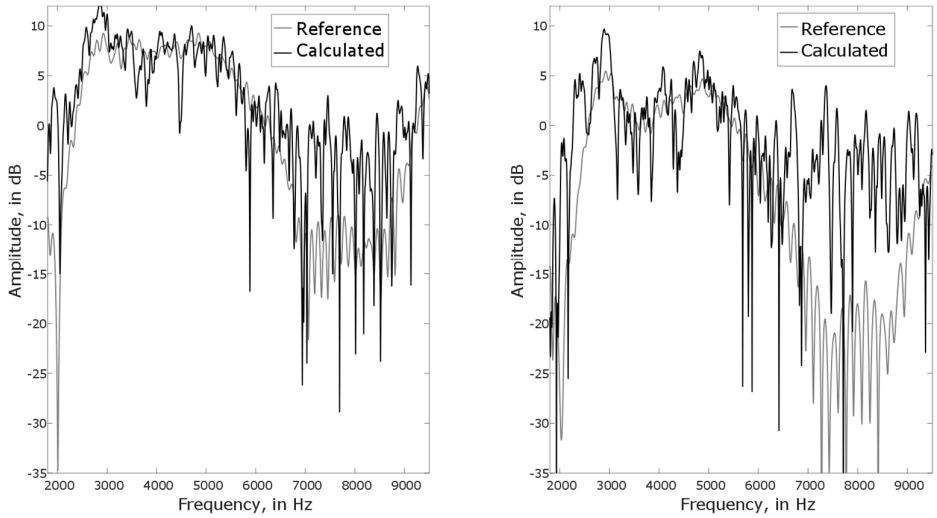


Fig. 5-24 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the left A-pillar (first point), car A.

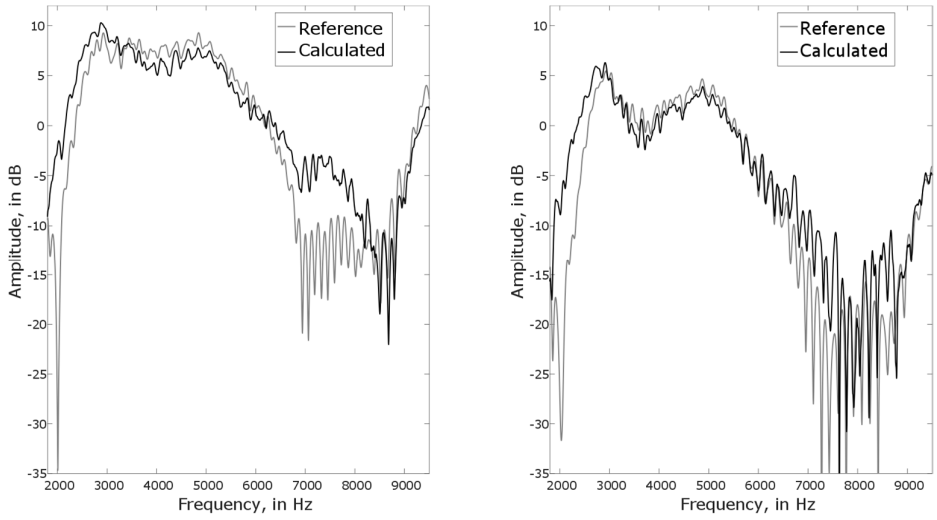


Fig. 5-25 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the left A-pillar (first point), car A

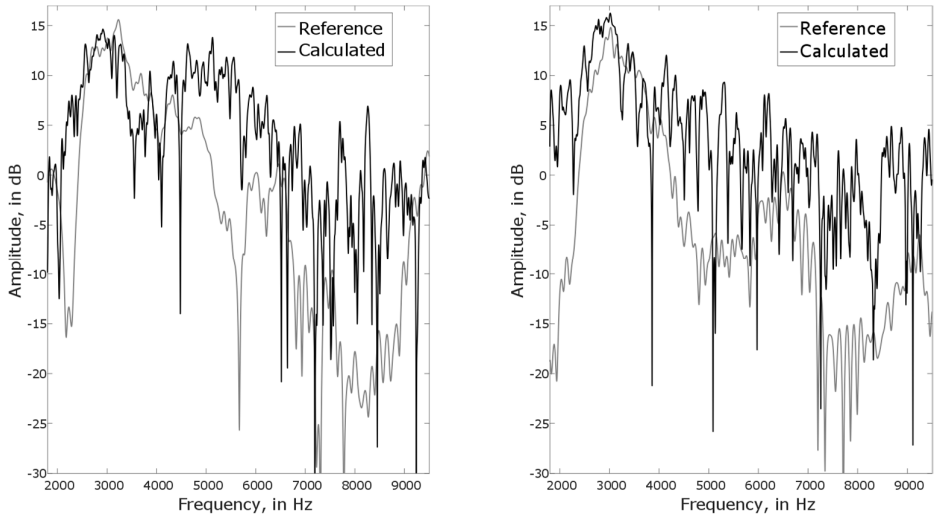


Fig. 5-26 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the left A-pillar (first point), car B.

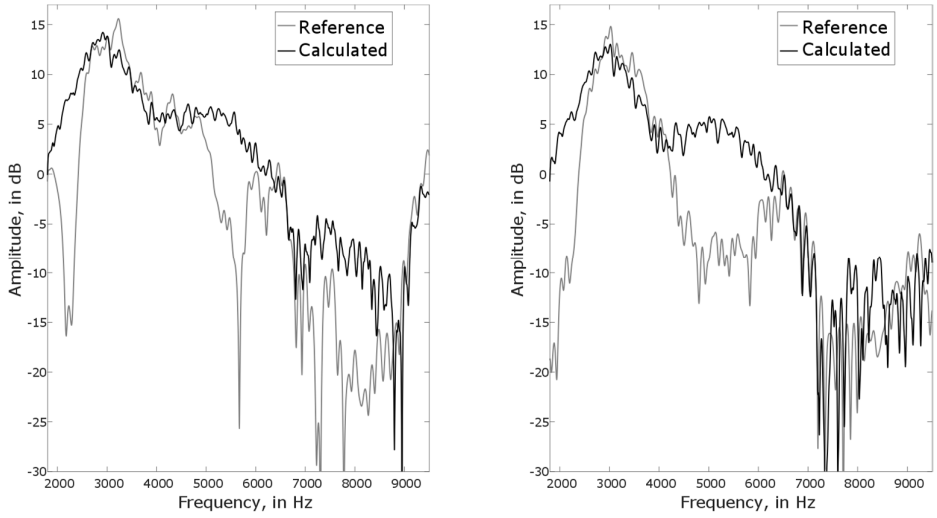


Fig. 5-27 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the left A-pillar (first point), car B.

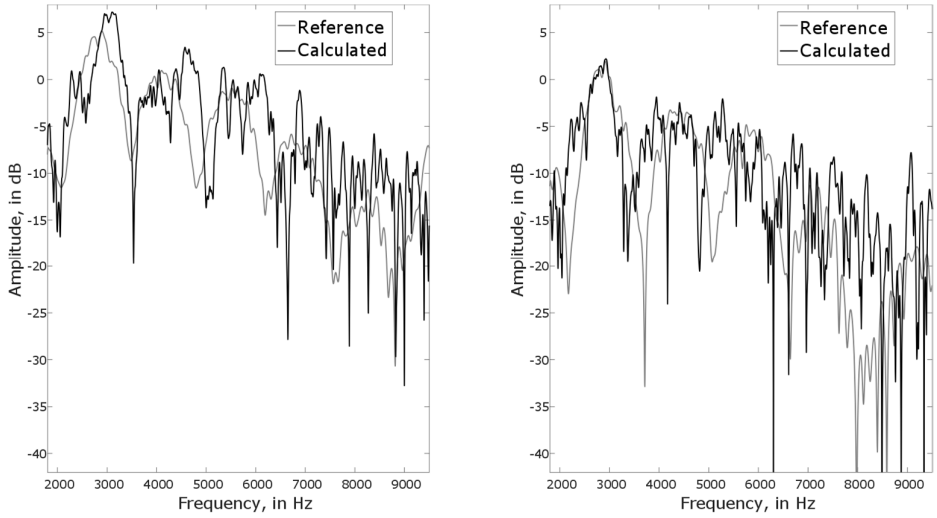


Fig. 5-28 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the left A-pillar (second point), car A.

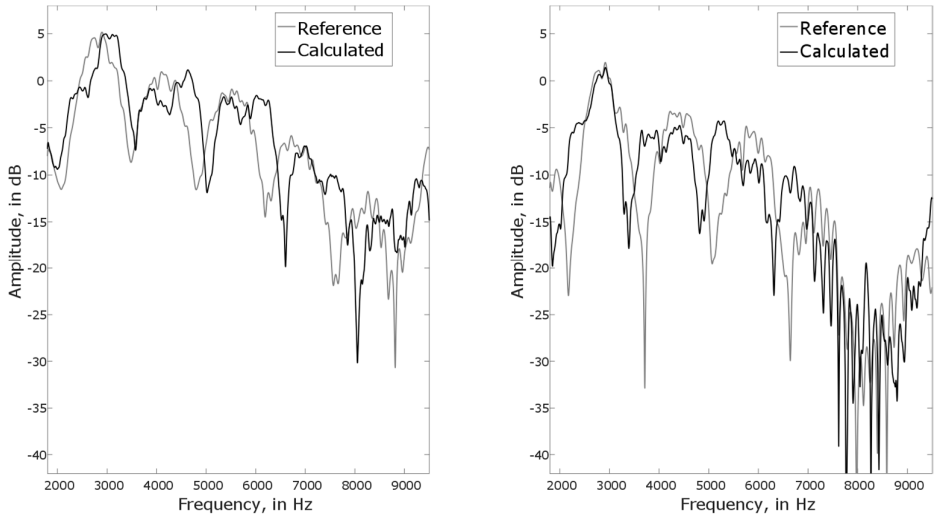


Fig. 5-29 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the left A-pillar (second point), car A.

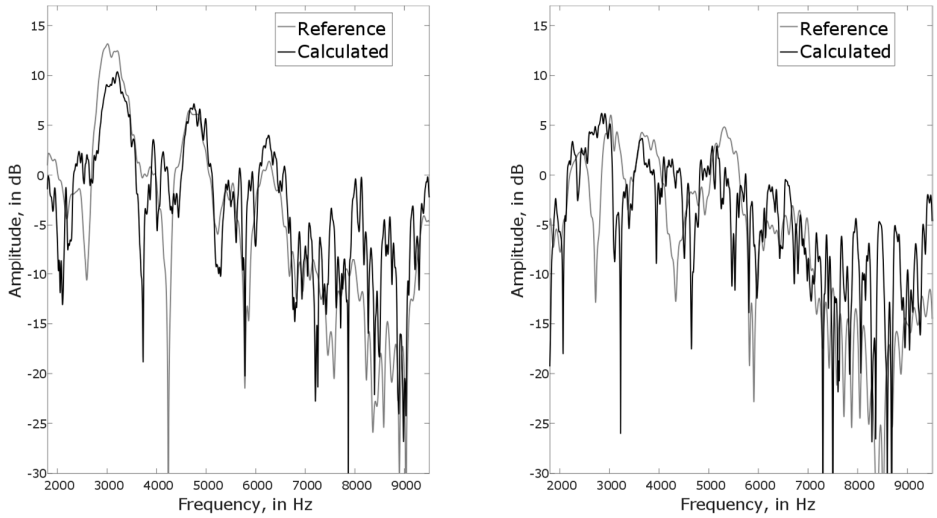


Fig. 5-30 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the left A-pillar (second point), car B.

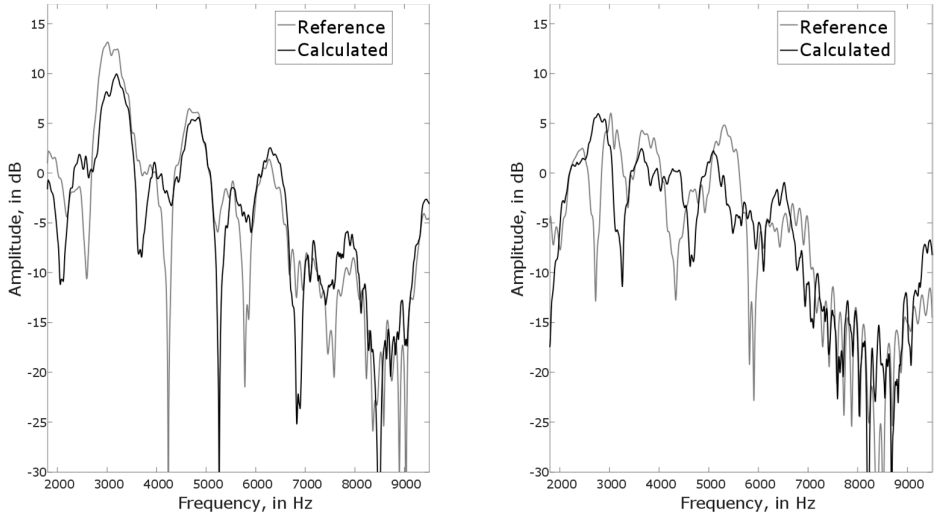


Fig. 5-31 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the left A-pillar (second point), car B.

Fig. 5-32 to Fig. 5-35 presents RTFs calculated for the wipers noise for the Car A and the Car B. Like for the left A-pillar (first point) strong early reflections from a dashboard and a windshield cause some differences in RTF shape. In Fig. 5-34 (right) is clearly visible effect of a direct path and a strong reflection interference pattern mismatch. Reference RTF has a notch filter at 3.5kHz, whereas calculated by the statistical model has it at 4kHz.
Fig. 5-36 to Fig. 5-47 show RTFs calculated by the model for the right A-pillar (two points) and the right mirror for the Car A and the Car B. It can be noticed that due to different source position (source moved towards right with respect to previous sources) now RTFs for the right ear is calculated more accurately than the left ear. It is obviously caused by the fact that the right ear is now the ipsilateral ear and the left ear – contralateral. A direct path arrives for the ipsilateral only, therefore RTFs can be calculated more accurately by the statistical model.

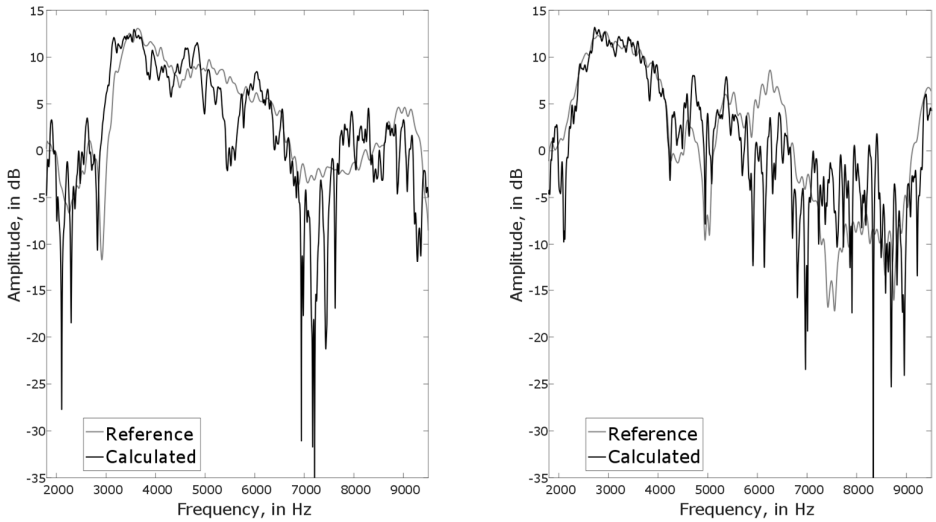


Fig. 5-32 RFTs calculated by the model with averaged parameters, left and right ear respectively, for wipers, car A.

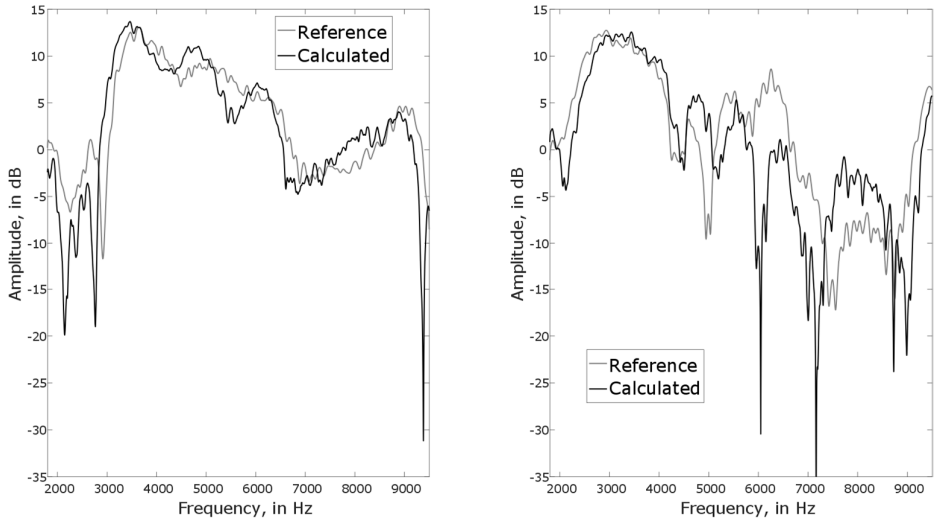


Fig. 5-33 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for wipers, car A.

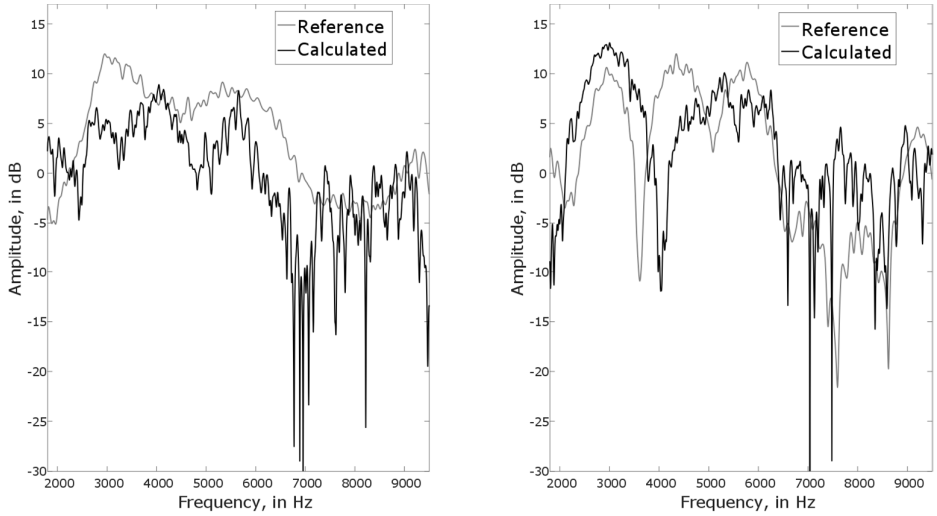


Fig. 5-34 RFTs calculated by the model with averaged parameters, left and right ear respectively, for wipers, car B.

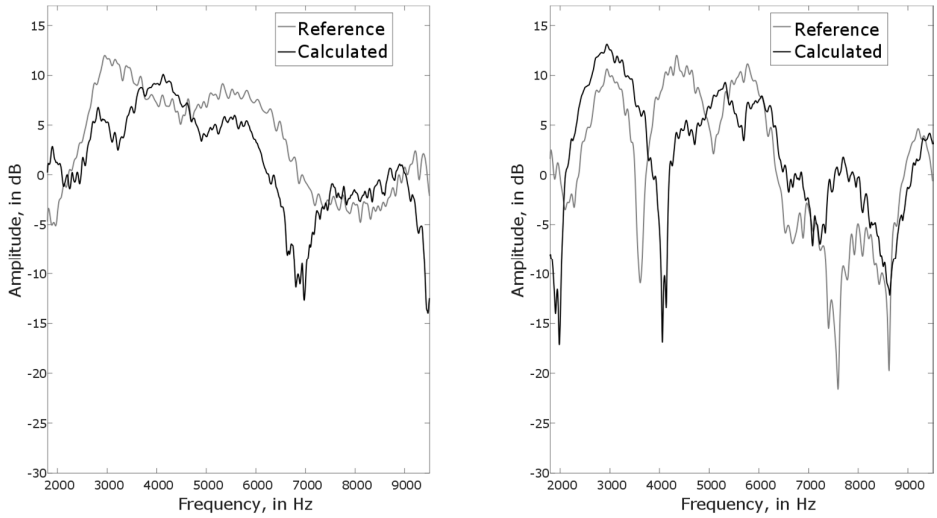


Fig. 5-35 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for wipers, car B.

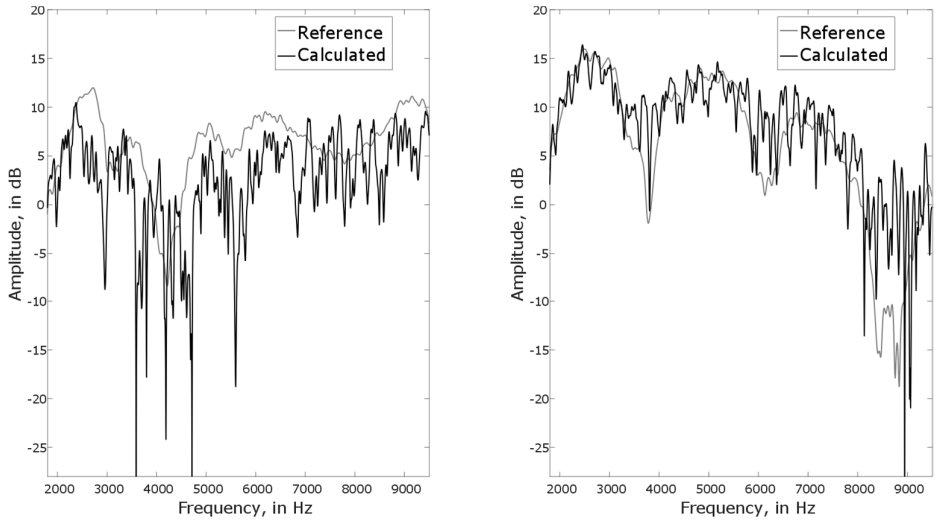


Fig. 5-36 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the right A-pillar (first point), car A.

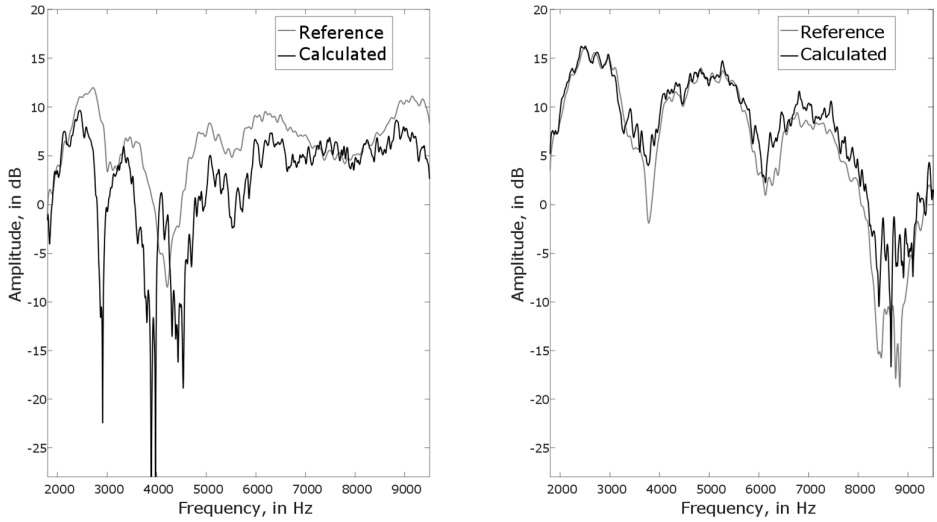


Fig. 5-37 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the right A-pillar (first point), car A.

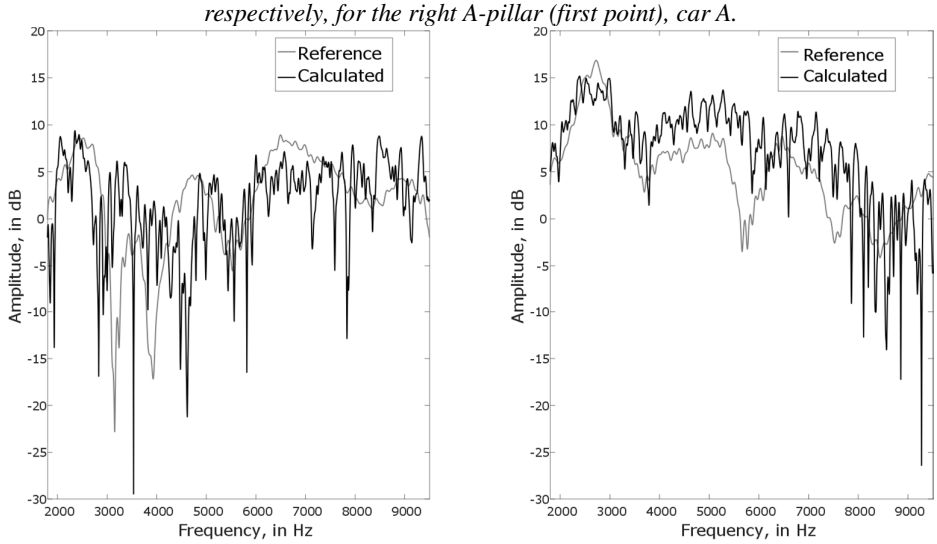


Fig. 5-38 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the right A-pillar (first point), car B.

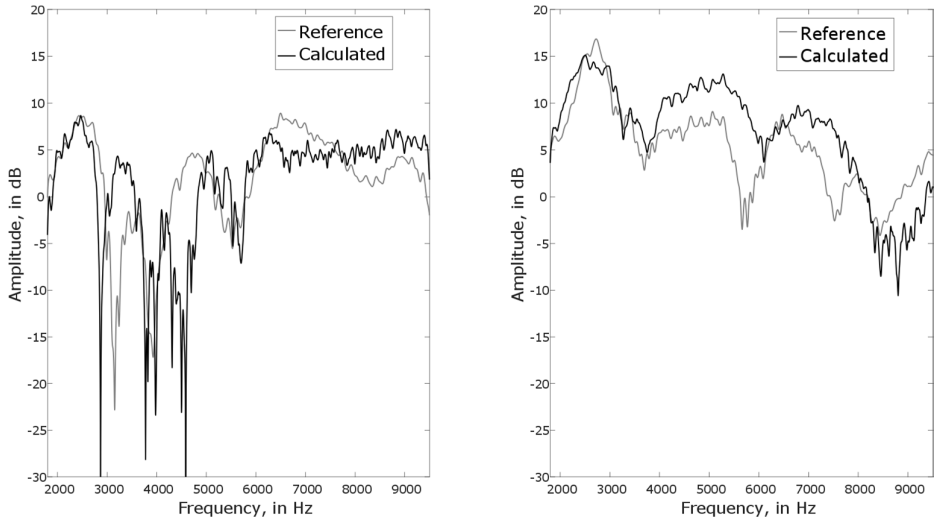


Fig. 5-39 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the right A-pillar (first point), car B.

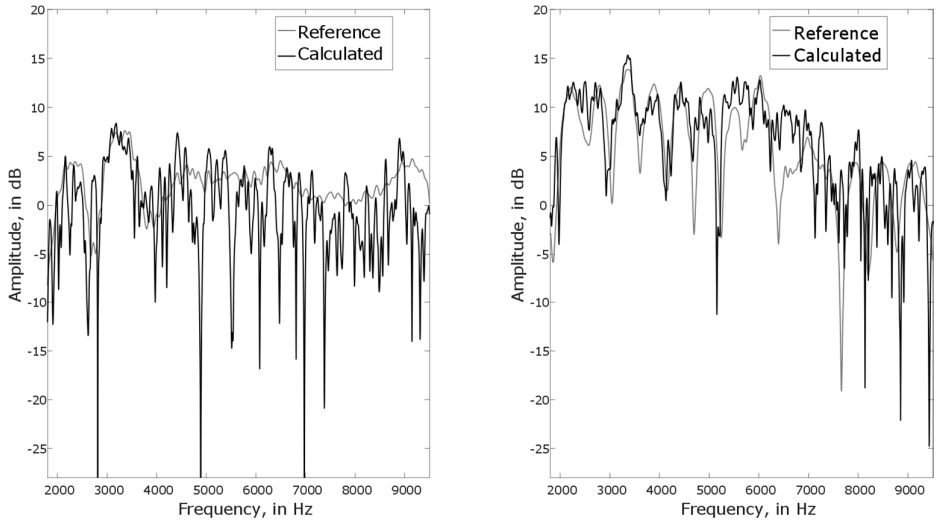


Fig. 5-40 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the right A-pillar (second point), car A.

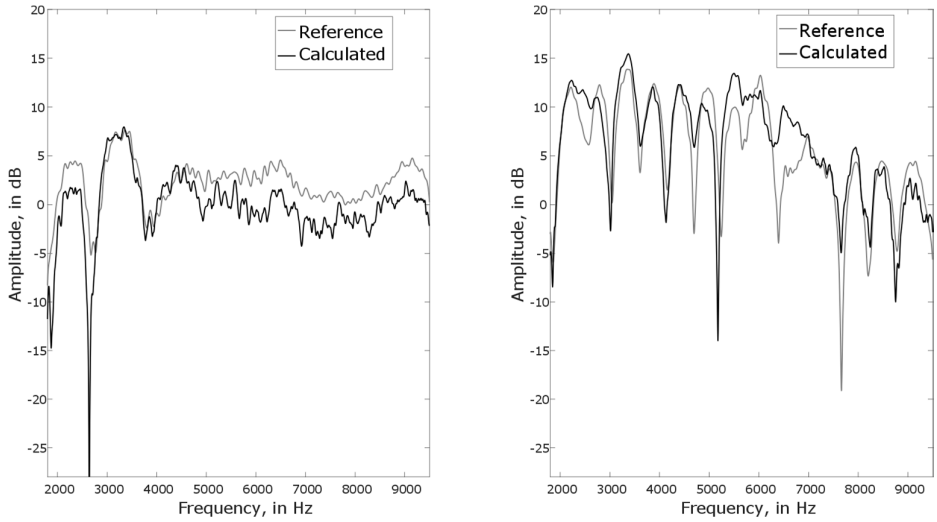


Fig. 5-41 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the right A-pillar (second point), car A.

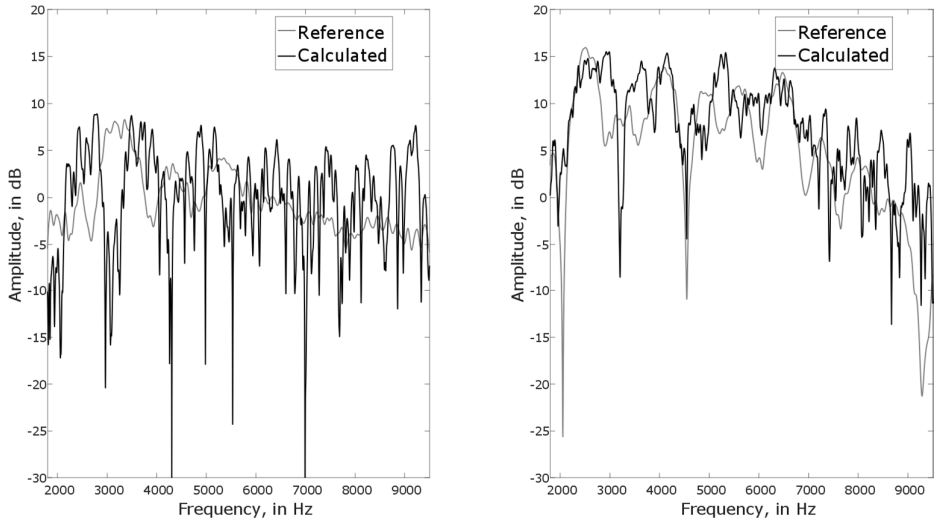


Fig. 5-42 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the right A-pillar (second point), car B.

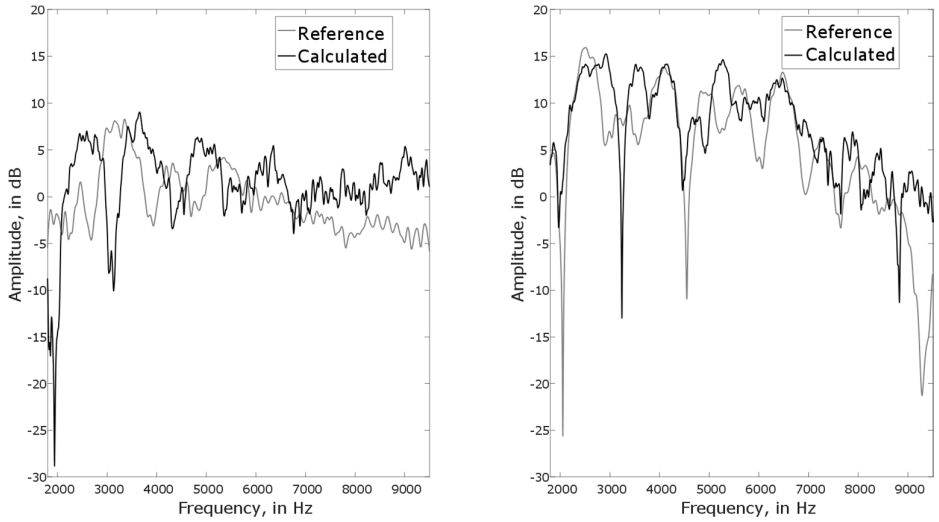


Fig. 5-43 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the right A-pillar (second point), car B.

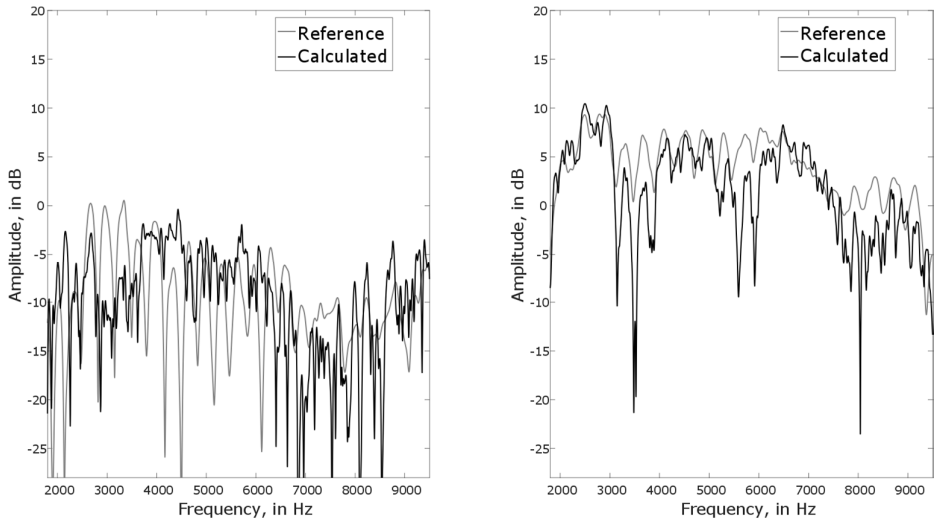


Fig. 5-44 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the right mirror, car B.

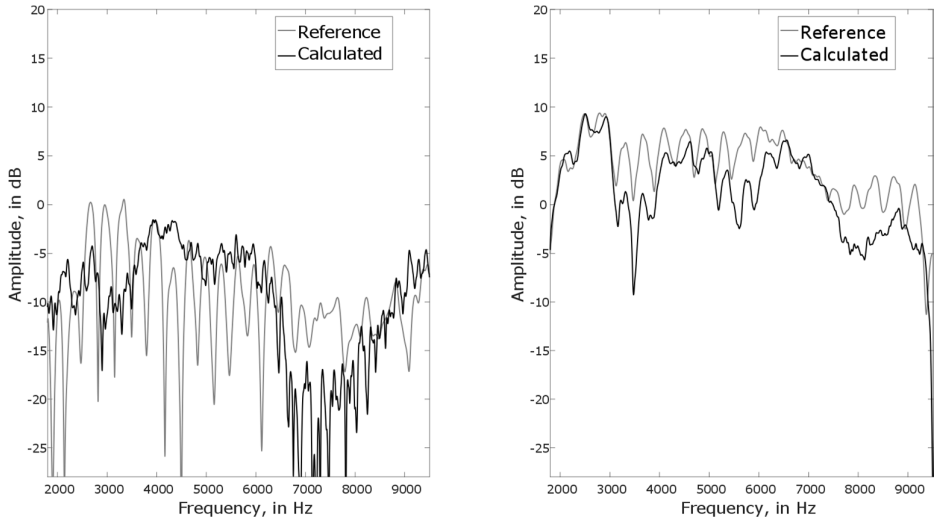


Fig. 5-45 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the right mirror, car B.

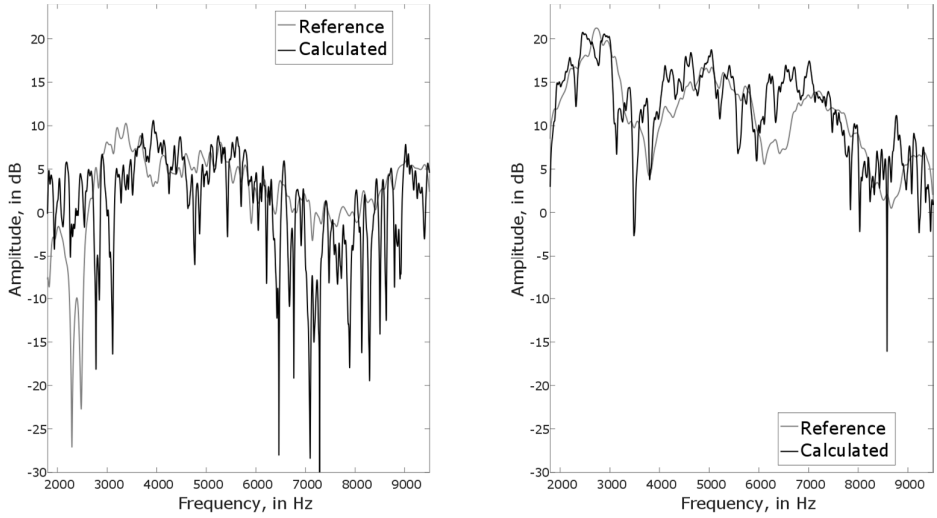


Fig. 5-46 RFTs calculated by the model with averaged parameters, left and right ear respectively, for the right mirror, car B.

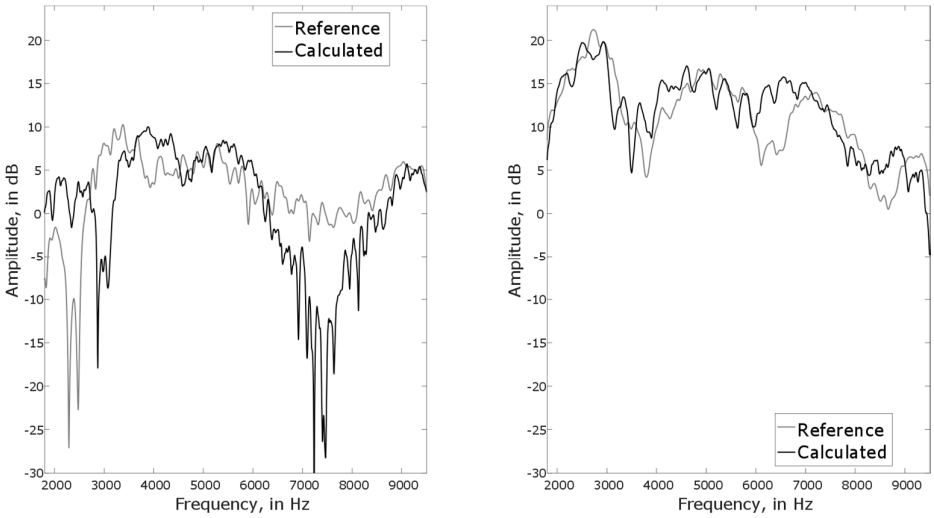


Fig. 5-47 RFTs calculated by the model with "tuned" parameters, left and right ear respectively, for the right mirror, car B.

5.5.1 Listening experiment

The most suitable method to validate the quality of RTFs generated by the statistical model is to perform a listening experiment. Twelve listeners without known hearing problems participated in the experiment (4 females and 8 males, ages 25–36). In the experiment, sounds were presented in pairs. Each pair consisted of a reference sound and a testing sound.. Tested sounds were wind noise sounds calculated using the statistical model. The reference sound was a wind noise sound calculated using TR simulation for the same conditions as the tested sound. List of presented sounds with content description is presented in Table 5-1.

Listeners were asked to rate any kind of perceived difference (it can be spectral or spatial difference) on a scale from 1 to 5, where 1 means no difference in sounds and 5 means that sounds are totally different. The order of appearance of the reference sound and the sound under investigation was randomized. The experiment was preceded by a training session. All sound were presented over AKG K550 headphones. Results are shown in Fig. 5-48.

Table 5-1 sound samples used in the listening experiment

Sound no.	Sound content
1	left mirror Car A
2	right mirror Car A
3	left mirror Car B
4	right mirror Car B
5	right A-pillar Car A
6	left A-pillar Car B
7	Left A-pillar Car A
8	Right A-pillar Car B
9	wipers Car A
10	left and right mirror + left and right A-pillars + wipers Car A
11	left and right mirror + left and right A-pillars + wipers Car B
12	right mirror Car A (reference Car B)
13	Left A-pillar Car A (reference Car B)
14	right A-pillar Car B (reference Car A)
15	left mirror Car A (reference Car B)
16	left and right mirror + left and right A-pillars + wipers Car B (ref. Car A)
17	left mirror Car A (statistical car)
18	right mirror Car A (statistical car)
19	left mirror Car B (statistical car)
20	right mirror Car B (statistical car)
21	right A-pillar Car A (statistical car)
22	left A-pillar Car B (statistical car)
23	Left A-pillar Car A (statistical car)
24	Right A-pillar Car B (statistical car)
25	left and right mirror + left and right A-pillars + wipers Car A (statistical car)
26	left and right mirror + left and right A-pillars + wipers Car B (statistical car)

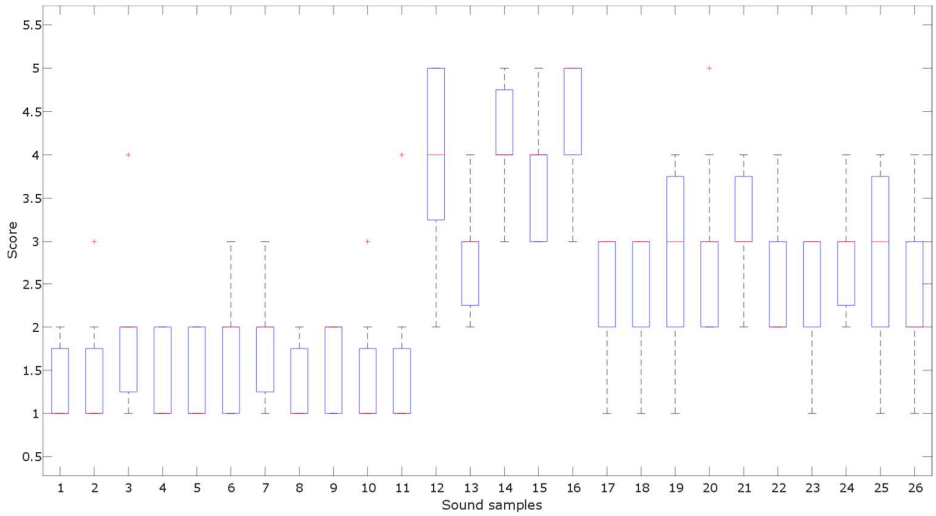


Fig. 5-48 Listening experiment results

Sounds 1-11 are mostly undistinguishable from the reference sounds. Differences between listeners responses for these sounds are statistically insignificant. When listeners notice a difference, it is usually cause by a different location of the source. Even though direction of arrival of the direct sound is known, strong early reflections appearing within 1ms after the direct sound disturb the precedence effect causing image shift. Fortunately these reflections usually arrive from almost the same direction, therefore reproduced sounds are only slightly different than the reference.

Sound 12-16 according to predictions are clearly different form the reference sounds. The purpose of presenting these sounds is to investigate significance of using correct HRTF and statistical model parameters. Due to different car sizes the same type of sources might appear even 15° away from each other. Majority of listeners perceived strong differences between test sounds and the reference. These listeners responses are in compliance with results achieved in Chapter 4.

The last sounds (17-26) were calculated using the "statistical car" model parameters. Listeners assess differences mostly from "slightly different" to "different". This result confirms the necessity of "tuning" the statistical model. Using model parameters of an average car only for some sources and cars can be at best "slightly different".

5.6 Conclusions

In this chapter a new statistical method for in-vehicle wind noise auralization was proposed. A mathematical description of the model and methods for parameters estimation were presented. The new model uses statistical distribution of RTFs parameters like DRR, ETL, EDT, RT60. A high accuracy database from eleven cars (RT simulated and measured) was prepared. Another important aspect of the model is treatment of early reflections. Number of early reflections used in the in the model varies from 3-5 depending on source location. Measurement points closer to reflecting surfaces (A-pillars, wipers) require more early reflections i.e. 5. For measurement points like mirror noise 3 reflections are sufficient.
Presented RTFs calculated by the model and performed listening experiment confirms that the new method can provide indistinguishable or only slightly different RTFs from measured or RT simulated RTFs.

Chapter 6

Conclusions and Future Work

In this thesis, the topic of in-vehicle auralization was covered. This research theme was tackled, in the document, discussing findings that could help in improving already existing auralization methods (Chapters 2, 3 and 4) and presenting, in Chapter 5, a new auralization approach. In this final conclusion, the most important results collected from each chapter are summarized below.

In Chapter 2 investigations on drivers' position on HRTFs are presented. Different azimuth and elevation angles were measured as well as source-receiver distances. It was found that presence of arms only slightly affects far-field HRTFs. On the contrary, as source gets closer to receiver (less than 1m), interaural cues have been found to change with distance due to presence of arms. In near field, contribution of arms increases with distance. At 70cm and 120cm distance reflections from arms are clearly noticeable, whereas they are less evident at 40cm. Such phenomenon is interesting for in-vehicle sound auralization since different source-receiver distances can occur. For example, in order to auralize left mirror and A-pillar wind noise, near field HRTFs should be used. For mirror and A-pillar on the right part of the car cabin (passenger position), using the HRTF database for 120cm is more suitable. For the ipsilateral ear influence of arms is very weak, whereas for the contralateral ear HRFT is strongly affected by the reflections from arms. Contralateral ear is always in acoustic shadow and, therefore, reflections from arms have higher impact on the HRTF.

Chapter 3 investigated perceptual aspects of RIR early reflections in car cabins. Two scenarios were presented: when only one source is present or with multiple sources masking the modified RIR. Considering wind noise source, the optimal number of reflections used in an auralization process is defined by the level of individual reflection as well as by the direction of arrival. For reflections appearing from a similar direction as a direct sound, the most important are those with a level higher than -10dB relative to a direct sound. Listeners are more sensitive to lateral reflections, where the threshold is approximate -14dB relative to a direct sound. When all wind noise sources are acting together, perception of early reflections in car cabins decreases and it becomes independent form direction of arrival of the reflection.

In Chapter 4, a human-oriented approach to sound sources localization accuracy in car cabins was investigated. Two listening experiments were carried out in order to examine perception of source small position differences. The first experiment, based on simulated and

measurement data, was aiming at investigating MAA of one wind noise source acting at a time. A data set of simulated sources evenly distributed on the windshield with 2° angle resolution within the range of 40° from the left to the right with respect to the receiver's position was used in the experiment. The receiver was placed at the centre of driver's head. Minimum Audible Angles were derived from a listening experiment for two vertical planes and the horizontal plane. MAA for the horizontal plane varies from 2° towards left to 4° towards right with respect to the head orientation. For the vertical plane at 0° azimuth (median plane) MAA is approximately 2° higher than in the horizontal plane. MAA for vertical plane is 2° lower when a source is moved slightly towards MAA for the horizontal plane. The second experiment used high accuracy measurement RIR data of a car interior. Scenarios with more than one sources acting at the same time were investigated. Introducing other sound sources while determining MAA changes final results. Results are significantly different compared to the first experiment. For the vertical plane adding other noise sources makes the value of MAA to increase from 2° to 6° for left direction (towards car cabin interior). Wipers noise MAA are slightly changed in horizontal plane, but significantly in the median plane. Range of MAA increased from 4° to 10°. The biggest influence of the presence of other noise sources can be observed for right mirror results. The range of MAA increased from 4° to 10° for the median plane and from 6° to 16° for the horizontal plane. This high difference is a consequence of different distance between the driver and left right mirrors. Left mirror generates louder noise at the drivers' position, therefore right mirror noise can be partially masked. Masking effect is good news for sound source localization accuracy problem. For only one source present at a time, the minimal range of MAA is, for instance in the horizontal plane, less than 4cm (left mirror). With wiper and right mirror noise added range of unnoticed source displacement raised to 8 cm. For the right mirror, requirements are bit less challenging: 6cm for the elevation. Masking has the highest impact on the right mirror noise localization accuracy, where maximal possible error increased from 6cm (median plane) to more than 15 cm. These values can be interpreted as the accuracy required for a source localization technique if its results are aimed at being exploited for auralization purposes.

Chapter 5 is mainly focused on the development of a new time-saving auralization method, tailored for car interior applications, which preserves sufficient fidelity in terms of sound reproduction. The proposed method aims to make it possible to auralize wind noise without any prior knowledge of the geometry of the car under investigation. This goal is achieved by using a statistical model of key Room Impulse Response (RIR) parameters and early reflections. Additionally, by exploiting sound source localization and separation techniques (e.g. time domain beamforming and Transfer Path Analysis) it becomes possible to virtually relocate sources and to perform what-if analyses. Currently, in order to "turn off" wind noise sources like side mirrors and evaluate the influence of this part to the overall noise, it is necessary to dismount it from the vehicle and to perform measurement in a wind tunnel again.

A mathematical description of the model is given with the method of combining RIR parameters. For the purpose of the statistical analysis, a big database of measurement and simulated RIR was generated. RIR were parametrized using RT, DRR, EDT, ETL. The method of data clustering was also included.

This Chapter presents additional interesting information on the acoustic behavior of car cabins when they are excited by wind noise sources. A statistical distribution of RIR parameters are presented and differences between clusters areas in car cabins are discussed and explained. In particular, differences in DRR and ETL are interesting from the statistical model point of view.

Correlations between in-vehicle clusters are presented. For some clusters, it is possible to calculate values of RIR parameters based on measurement performed in different cluster. These information can be used in order reduce number of RIR measurement points for auralization.

To generalize the statistical model for different type of noise and source positions, the early reflection pattern can be possibly substituted with the Image Source Method. Of source, geometries of car cabins are not well suited for this method. Moreover, the big advantage of proposed method is auralization without a cavity model at all. However, simplified model of a car, like presented in Fig. 4-1, can be possible to implement with the Image Source Method. One model could be sufficient to predict early reflection behavior. Thorough investigations on influence of this simplification should be done.

Used measurement point clustering method for statistical model allows straightforward comparison of RIR parameters between cars. However, in-cluster spread of parameter values is high. Therefore, other clustering methods can be investigated resulting lower in-cluster standard deviation.

References

[1] M. Vorländer, Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality, Springer Science & Business Media, 2007.

[2] H. Wallach, E. B. Newman and M. R. Rosenzweig, “A precedence effect in sound localization,” *The Journal of the Acoustical Society of America,* vol. 21, no. 4, pp. 468-468, 1949.

[3] M. Barron, “The subjective effects of first reflections in concert halls - the need for lateral reflections,” *Journal of sound and vibration,* vol. 15, pp. 475-494, 1971.

[4] P. Rubak, “Coloration in room impulse responses,” in *Proc. Joint Baltic-Nordic Acoustics Meeting*, 2004.

[5] J. Blauert, Spatial hearing: the psychophysics of human sound localization, MIT press, 1997.

[6] C. F. Eyring, “Reverberation time in "dead" rooms,” *The Journal of the Acoustical Society of America,* vol. 1, pp. 217-241, 1930.

[7] A. Krokstad, S. Strom and S. Sørsdal, “Calculating the acoustical room response by the use of a ray tracing technique,” *Journal of Sound and Vibration,* vol. 8, pp. 118-125, 1968.

[8] M. Aretz, Combined wave and ray based room acoustic simulations of small rooms, vol. 12, Logos Verlag Berlin GmbH, 2012.

[9] J.-J. Embrechts, “Broad spectrum diffusion model for room acoustics ray-tracing algorithms,” *The Journal of the Acoustical Society of America,* vol. 107, pp. 2068-2081, 2000.

[10] M. Vorländer, “Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm,” *The Journal of the Acoustical Society of America,* vol. 86, pp. 172-178, 1989.

[11] K. H. Kuttruff, “Auralization of impulse responses modeled on the basis of ray-tracing results,” *Journal of the Audio Engineering Society,* vol. 41, pp. 876-880, 1993.

[12] A. M. Ondet and J. L. Barbry, “Modeling of sound propagation in fitted workshops using ray tracing,” *The Journal of the Acoustical Society of America,* vol. 85, pp. 787-796, 1989.

[13] M. R. Schroeder, “Frequency-Correlation Functions of Frequency Responses in Rooms,” *The Journal of the Acoustical Society of America,* vol. 34, pp. 1819-1823, 1962.

[14] J. A. Moorer, “About this reverberation business,” *Computer music journal,* pp. 13-28, 1979.

[15] J.-D. Polack, “La transmission de lnergie sonore dans les salles,” 1988.

[16] H. Kuttruff, Room acoustics, Crc Press, 2016.

[17] P. Rubak and L. G. Johansen, “Artificial reverberation based on a pseudo-random impulse response II,” in *Audio Engineering Society Convention 106*, 1999.

[18] T. Hidaka, Y. Yamada and T. Nakagawa, “A new definition of boundary point between early reflections and late reverberation in room impulse responses,” *Journal of the Acoustical Society of America,* vol. 122, pp. 326-332, 2007.

[19] M. R. Schroeder, “Response to Comments on New Method of Measuring Reverberation Time [PW Smith, Jr., J. Acoust. Soc. Am. 38, 359 (L)(1965)],” *The Journal of the Acoustical Society of America,* vol. 38, pp. 359-361, 1965.

[20] J. S. Abel and P. Huang, “A simple, robust measure of reverberation echo density,” in *Audio Engineering Society Convention 121*, 2006.

[21] G. Defrance, L. Daudet and J.-D. Polack, “Using matching pursuit for estimating mixing time within room impulse responses,” *Acta Acustica united with Acustica,* vol. 95, pp. 1071-1081, 2009.

[22] R. Stewart and M. Sandler, “Statistical measures of early reflections of room impulse responses,” in *Proc. of the 10th int. conference on digital audio effects (DAFx-07), Bordeaux, France*, 2007.

[23] A. Lindau, L. Kosanke and S. Weinzierl, “Perceptual evaluation of model-and signal-based predictors of the mixing time in binaural room impulse responses,” *Journal of the Audio Engineering Society,* vol. 60, pp. 887-898, 2012.

[24] E. A. P. Habets, *Single- and multi-microphone speech dereverberation using spectral enhancement,* vol. 68, 2007.

[25] C. S. J. Doire, M. Brookes, P. A. Naylor, D. Betts, C. M. Hicks, M. A. Dmour and S. H. Jensen, “Single-channel blind estimation of reverberation parameters,” in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, 2015.

[26] M. R. Schroeder, “Integrated-impulse method measuring sound decay without using impulses,” *The Journal of the Acoustical Society of America,* vol. 66, pp. 497-500, 1979.

[27] V. L. Jordan, “Room acoustics and architectural acoustics development in recent years,” *Applied Acoustics,* vol. 2, pp. 59-81, 1969.

[28] P. Zahorik, D. S. Brungart and A. W. Bronkhorst, “Auditory distance perception in humans: A summary of past and present research,” *Acta Acustica united with Acustica,* vol. 91, pp. 409-420, 2005.

[29] HEAD acoustics GmbH, [Online]. Available: https://www.head-

acoustics.de/eng/nvh_consulting_automotive_applications_wind_noise.htm. [Accessed 2017].

[30] G. Cerrato, "Automotive sound quality--powertrain, road and wind noise," *Sound \& vibration,* vol. 43, pp. 16-24, 2009.

[31] M. Cho, H. G. Kim, C. Oh, K. D. Ih, A. Khondge, F. Mendonca, J. Lim, E.-S. Choi, B. Ganty and R. HALLEZ10, "Benchmark study of numerical solvers for the prediction of interior noise transmission excited by A-pillar vortex," in *Inter-Noise Conference Proceedings*, 2014.

[32] P. Majdak, Y. Iwaya, T. Carpentier, R. Nicol, M. Parmentier, A. Roginska, Y. Suzuki, K. Watanabe, H. Wierstorf, H. Ziegelwanger and others, "Spatially oriented format for acoustics: A data exchange format representing head-related transfer functions," in *Audio Engineering Society Convention 134*, 2013.

[33] "http://www.kfs.oeaw.ac.at," [Online]. [Accessed 2 May 2017].

[34] B. Gardner, K. Martin and others, "HRTF measurements of a KEMAR dummy-head microphone," *Massachusetts Institute of Technology,* vol. 280, pp. 1-7, 1994.

[35] B. Bernschutz, "A spherical far field HRIR/HRTF compilation of the Neumann KU 100," in *Proceedings of the 40th Italian (AIA) Annual Conference on Acoustics and the 39th German Annual Conference on Acoustics (DAGA) Conference on Acoustics*, 2013.

[36] P. F. F. Hoffmann, "Characteristics of head-related transfer functions," 2008.

[37] S. Bech, "Timbral aspects of reproduced sound in small rooms. I," *Acoustical Society of America. Journal,* vol. 97, pp. 1717-1726, 1995.

[38] S. Bech, "Timbral aspects of reproduced sound in small rooms. II," *Acoustical Society of America. Journal,* vol. 99, pp. 3539-3550, 1996.

[39] S. E. Olive and F. E. Toole, "The detection of reflections in typical rooms," *Journal of the Audio Engineering Society,* vol. 37, pp. 539-553, 1989.

[40] D. R. Begault, "Audible and inaudible early reflections: thresholds for auralization system design," in *Audio Engineering Society Convention 100*, 1996.

[41] D. R. Begault, B. U. McClain and M. R. Anderson, "Early reflection thresholds for anechoic and reverberant stimuli within a 3-D sound display," in *Proc. 18th Int. Congress on Acoust.(ICA04), Kyoto, Japan*, 2004.

[42] P. Zahorik, "Direct-to-reverberant energy ratio sensitivity," *The Journal of the Acoustical Society of America,* vol. 112, pp. 2110-2117, 2002.

[43] E. Larsen, N. Iyer, C. R. Lansing and A. S. Feng, "On the minimum audible difference in direct-to-reverberant energy ratioa)," *The Journal of the Acoustical Society of America,* vol. 124, pp. 450-461, 2008.

[44] P. M. Zurek, "Measurements of binaural echo suppression," *The Journal of the Acoustical Society of America,* vol. 66, pp. 1750-1757, 1979.

[45] A. M. Salomons, Coloration and binaural decoloration of sound due to reflections, TU Delft, Delft University of Technology, 1995.

[46] D. R. Begault, "Perceptual effects of synthetic reverberation on three-dimensional audio systems," *Journal of the Audio Engineering Society,* vol. 40, pp. 895-904, 1992.

[47] B. Bernschütz, "A spherical far field HRIR/HRTF compilation of the Neumann KU 100," in *Proceedings of the 40th Italian (AIA) Annual Conference on Acoustics and the 39th German Annual Conference on Acoustics (DAGA) Conference on Acoustics*, 2013.

[48] P. Castellini and M. Martarelli, "Acoustic beamforming: analysis of uncertainty and metrological performances," *Mechanical Systems and Signal Processing,* vol. 22, pp. 672-692, 2008.

[49] A. W. Mills, "On the minimum audible angle," *The Journal of the Acoustical Society of America,* vol. 30, no. 4, pp. 237-246, 1958.

[50] J. Blauert, *Spatial hearing,* Cambridge, Ma: MIT Press, 1997.

[51] B. Moore, An Introduction to the Psychology of Hearing, Brill, 2012.

[52] M. Barron, "The subjective effects of first reflections in concert halls - the need for lateral reflections," *Journal of sound and vibration,* vol. 15, no. 4, pp. 475-494, 1971.

[53] M. Lofdahl, A. Nykanen and R. Johnsson, "Assessment of changes in automotive sounds caused by displacements of source and listening positions," *Noise Control Engineering Journal,* vol. 60, no. 3, pp. 283-292, 2012.

[54] K. Kotorynski, "Digital binaural/stereo conversion and crosstalk cancelling," *Audio Engineering Society Convention 89,* 1990.

[55] P. Naylor and N. D. Gaubitch, Speech dereverberation, Springer Science & Business Media, 2010.

[56] B. W. Gillespie, H. S. Malvar and D. A. F. Florêncio, "Speech dereverberation via maximum-kurtosis subband adaptive filtering," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, 2001.

[57] M. R. Schroeder, "Statistical parameters of the frequency response curves of large rooms," *Journal of the Audio Engineering Society,* vol. 35, pp. 299-306, 1987.

[58] J. Eaton, N. D. Gaubitch and P. A. Naylor, "Noise-robust reverberation time estimation using spectral decay distributions with reduced computational cost," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013.

[59] J.-M. Jot, L. Cerveau and O. Warusfel, "Analysis and synthesis of room reverberation based on a statistical time-frequency model," in *Audio Engineering Society Convention 103*, 1997.

[60] C. Doire, M. Brookes, P. A. Naylor, C. M. Hicks, D. Betts, M. A. Dmour and S. H. Jensen, “Single-channel enhancement of speech corrupted by reverberation and noise,” *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP),* vol. 25, 2017.

[61] C. Colangeli, P. Chiariotti, G. Battista, P. Castellini and K. Janssens, “CLUSTERING INVERSE BEAMFORMING FOR INTERIOR SOUND SOURCE LOCALIZATION: APPLICATION TO A CAR CABIN MOCK-UP”.

[62] J. Tabor and P. Spurek, “Cross-entropy clustering,” *Pattern Recognition,* vol. 47, pp. 3046-3059, 2014.