

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

7,000

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Perspective Chapter: Modern Acquisition of Personalised Head-Related Transfer Functions – An Overview

*Katharina Pollack, Wolfgang Kreuzer and Piotr Majdak*

## Abstract

Head-related transfer functions (HRTFs) describe the spatial filtering of acoustic signals by a listener's anatomy. With the increase of computational power, HRTFs are nowadays more and more used for the spatialised headphone playback of 3D sounds, thus enabling personalised binaural audio playback. HRTFs are traditionally measured acoustically and various measurement systems have been set up worldwide. Despite the trend to develop more user-friendly systems and as an alternative to the most expensive and rather elaborate measurements, HRTFs can also be numerically calculated, provided an accurate representation of the 3D geometry of head and ears exists. While under optimal conditions, it is possible to generate said 3D geometries even from 2D photos of a listener, the geometry acquisition is still a subject of research. In this chapter, we review the requirements and state-of-the-art methods for obtaining personalised HRTFs, focusing on the recent advances in numerical HRTF calculation.

**Keywords:** head-related transfer functions, spatial hearing, acoustic measurement, numerical calculation, localisation

## 1. Introduction

Head-related transfer functions (HRTFs) describe the filtering of the acoustic field produced by a sound source arriving at the listener's ear. The filtering is the effect of the interaction of the sound field with the listener's anatomy and has various properties. First, the incoming sound wave arrives at the ipsilateral pinna, i.e., the ear closer to the sound source, and then at the contralateral ear, i.e., the ear away from the sound source. This time difference between ipsilateral and contralateral ear is usually described as the interaural time difference (ITD). Second, larger anatomical structures, i.e., torso, shoulders and head, affect frequencies up to 3 kHz in a comparatively trivial way. As the listener's torso and head shadow the sound wave arriving at the contralateral ear, interaural level differences (ILDs) arise. Third, the incoming sound is filtered in a complex way by the shape of the listener's pinnae. These monaural time-frequency-filtering effects become especially important for higher frequency regions (above approximately 4 kHz) and sound directions inducing the same ITDs and ILDs [1–6]. Humans have learned to interpret this acoustic filtering to span an auditory space as an internal model of

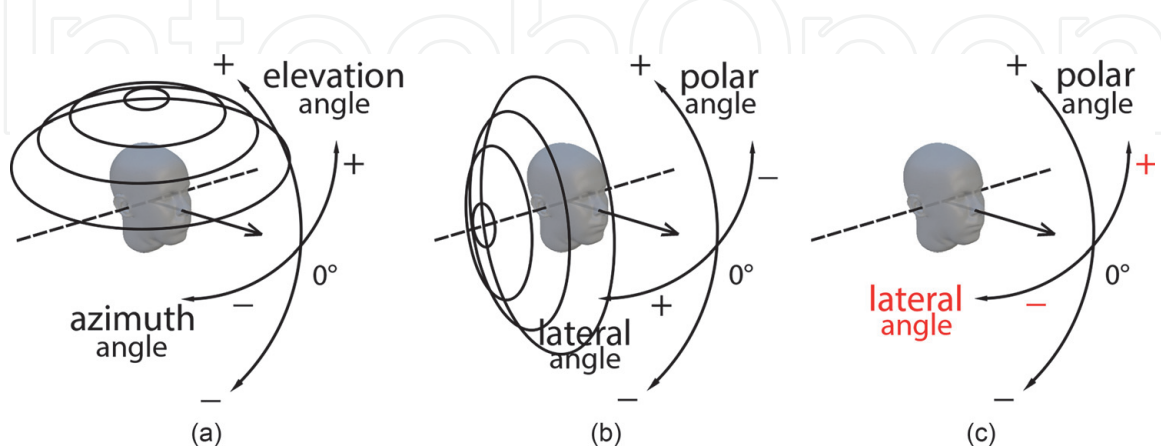
their natural environment [7]. Because the pinna shape is unique for every person, HRTFs are considered listener-specific [8–10], similar to a fingerprint [1–6]. With an individually fitted HRTF dataset, it is possible for a person to perceive sounds (in a virtual environment) via headphones as if the sounds would originate from their (physical) position around the listener.

Both interaural and monaural features for a single sound direction can be represented by a binaural HRTF pair [11]. In signal processing terms, a binaural HRTF pair can be described as

$$\begin{aligned} \text{HRTF}_L(\mathbf{x}^*, f, s) &= \frac{p_L(\mathbf{x}^*, f, s)}{p_0(0, f)} \\ \text{HRTF}_R(\mathbf{x}^*, f, s) &= \frac{p_R(\mathbf{x}^*, f, s)}{p_0(0, f)} \end{aligned} \quad (1)$$

where  $p_L$  and  $p_R$  describe the sound pressure at a position inside the left and right ear, respectively (typically the entrance of the left and right ear canal or a position close to the eardrum),  $\mathbf{x}^*$  describes the sound-source position (i.e., distance and direction),  $f$  describes the frequency and  $s$  the listener's geometry, emphasising the listener-specificity of HRTFs.  $p_0$  describes the reference sound pressure, which is usually the pressure measured at the position of the midpoint of right and left ear *without* the head being present.

There are several options to set a specific coordinate system to systematically describe directions for HRTFs. From the physical perspective, the *spherical* coordinate system is a natural choice; in that case, the origin of the system is placed inside the listener's head at the midpoint between left and right ear and the direction is described by azimuth and elevation angles, see **Figure 1a**. In this system, one can intuitively define the two main planes: The eye-level horizontal plane, i.e., all directions with the elevation angle of zero, and the median plane, i.e., all directions with the azimuth angle of zero. The eye-level horizontal plane is also called Frankfurt plane and can be anatomically defined as the plane connecting the lowest part of the listener's orbital cavity and the highest part of the bony ear canal (meatus acusticus externus osseus). This spherical coordinate system resembles a *geodesic* representation widely used in physics, with the poles located at the top and bottom. An alternative system that is more relevant from the auditory perspective is given by the *interaural-*



**Figure 1.** Coordinate systems typically used in the HRTF acquisition and representation. The dashed line represents the interaural axis, and the arrow represents the viewing direction. (a) Spherical coordinate system with the azimuth and elevation angles. (b) Simple interaural-polar coordinate system with the lateral and polar angles obtained by rotation the poles of the spherical system. (c) Modified interaural-polar coordinate system with the lateral and polar angles corresponding to the azimuth angle in the horizontal plane and the elevation angle in the median plane.

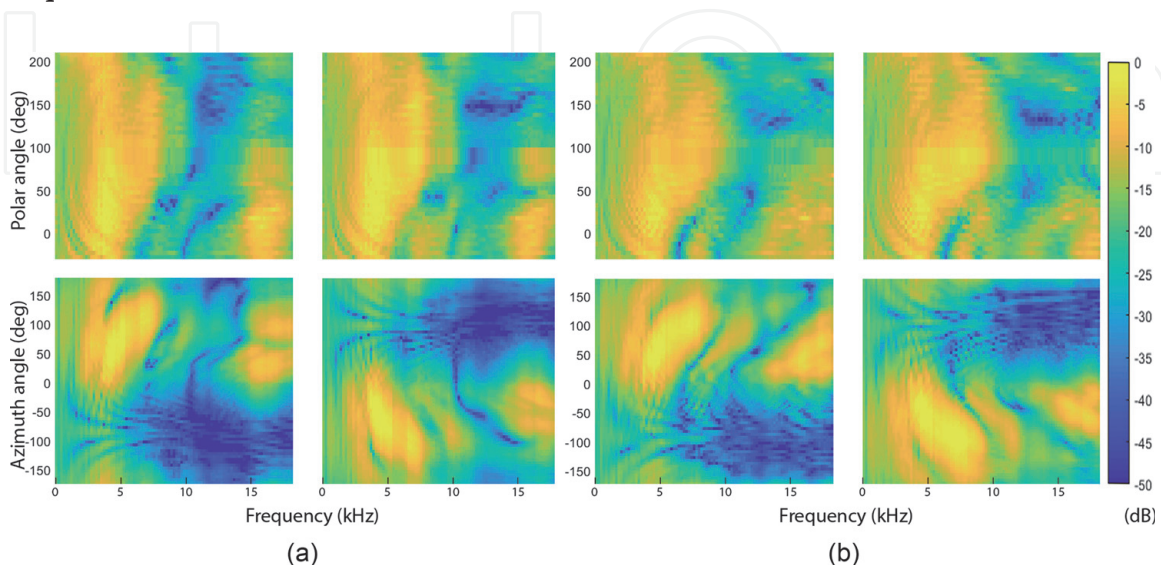
*polar* coordinate system. This system is shown in **Figure 1b** and can be constructed by rotating the poles of the spherical system to the interaural axis, i.e., the axis connecting the two ears. A sound direction is then described by the lateral angles (along the horizontal plane) and polar angles (along the median plane). The poles are then located on the left and right sides of the listener. This simple interaural-polar coordinate system was used in various psychoacoustic studies, e.g., [12, 13], and has the disadvantage that the lateral angle does not correspond to the azimuth angle.

**Figure 1c** shows the *modified* version of the interaural-polar coordinate system, which does not have this disadvantage. Here, the sign of the lateral angle is flipped, i.e., in the coordinate system, the positive lateral angles are used for sounds located on the left side of the listener. This transformation to a left-handed coordinate system has the advantage of having the lateral angle corresponding to the azimuth angle for all sources placed in the horizontal plane, and the polar angle corresponding to the elevation angle for all sources placed in the median plane. Thus, the modified interaural-polar coordinate system offers a better link between the psychoacoustic research and audio engineering. In that system, the lateral angle ranges from  $-90^\circ$  (right ear) over  $0^\circ$  (front) to  $90^\circ$  (left ear), and the polar angle ranges from  $-90^\circ$  (bottom) over  $0^\circ$  (front) and  $90^\circ$  (up) to  $180^\circ$  (back) and  $270^\circ$  (bottom again).

The understanding of these coordinate systems is important because state-of-the-art acquisitions and representations of HRTFs utilise those systems. For example, **Figure 2** shows HRTFs along the Frankfurt and the median plane. These various coordinate systems are used in HRTF visualisation, in various HRTF-related software packages such as the SOFA toolbox [15], and in auditory modelling, e.g., the Auditory Modelling Toolbox (AMT) [16, 17].

HRTF acquisition can be classified into three categories: acoustic measurement, numerical calculation, and personalisation [18].

The acoustic measurement is traditionally designed as the measurement of the impulse response between source and receiver in an anechoic or semianechoic chamber, describing the transmission path from a sound source to the ear [11, 19]. A comprehensive review of the established state-of-the-art acoustic techniques to measure HRTFs can be found in [20]. Thus, in this chapter, Section 3, we only briefly provide an overview of the traditional acoustic HRTF measurement approaches, highlight some of their differences and new trends and focus on the requirements for the acoustic measurement.



**Figure 2.** HRTF magnitude spectra for the listeners (a) NH236 and (b) NH257, both from the ARI database [14]. Top: Spectra along the median plane. Bottom: Spectra along the eye-level horizontal plane. 0 dB corresponds to the maximum magnitude in each panel.



Numerical HRTF calculation simulates the acoustic measurement by considering a 3D representation of the listener's geometry and the positions of multiple external sound sources, for which the generated sound pressure at the entrance of the ear canal is calculated. This technique has become more popular and is the main focus of this chapter. To this end, in Section 4, we provide an overview of the principles of various numerical calculation approaches including a comparison of the mentioned methods.

Personalisation of HRTFs describes the process of adapting an existing set of generic data guided by listener-specific information, either with the help of objective or subjective personalisation method. The objective personalisation has been approached from two different domains: the geometric domain, in which listener-specific anthropometric data are measured and used to personalise a generic geometric model from which HRTFs are then simulated; or the spectral domain, in which a generic HRTF set is directly personalised based on listener-specific information. Examples for personalisation approaches include utilising frequency scaling [21], parametric modelling of peaks and notches [22], active shape modelling (ASM) [23], principal component analysis (PCA) in both geometric [24] and spectral domains [25–29], multiple regression analysis [30], independent component analysis (ICA) [31], large deformation diffeomorphic metric mapping (LDDMM) [25, 32], local neighbourhood mapping [33], neural networks [34–41] and linear combination of HRTFs [42]. Despite many efforts worldwide [43–46], the link between the morphology and HRTFs is not fully understood yet, mostly because of the high dimensionality of the problem. Most recent tools for studying that link are rooted in aligning high-resolution pinna representations to target representations facilitated with parametric pinna models [47, 48].

In the subjective personalisation, listeners are confronted with several sets of HRTFs and an algorithm (usually based on the evaluation of localisation errors, i.e., the difference between perceived and actual sound-source location) adapts the HRTF sets aiming at converging at listener-specific HRTFs [9, 49]. For an educated guess for the initial sets, anthropometric data can be used to pre-scale the HRTF sets, or the HRTF sets can be pre-selected via psychoacoustic models [50]. Clustering of the HRTF sets can further improve the relevance and reduce the duration of the personalisation procedure [49, 51].

All these methods aim at providing a specific quality in terms of acoustic and psychoacoustic properties. In the following section, we describe the acoustic properties and psychoacoustic requirements for human HRTFs, both of which lay the base for HRTF acquisition. Then, we briefly describe the most important requirements for the acoustic HRTF measurement, complementing the work of Li and Peissig [20]. Finally, we describe approaches for numeric HRTF calculation in greater detail.

## **2. Head-related transfer functions: acoustic properties and psychoacoustic requirements**

In this section, we describe the acoustic properties of HRTFs and relate them to psychophysical properties of human hearing with the goal to derive the minimum requirements for sufficiently accurate HRTF acquisition by means of perception. We analyse spectral, temporal and spatial aspects of HRTFs and consider contributions of distinct parts of the human body to these aspects.

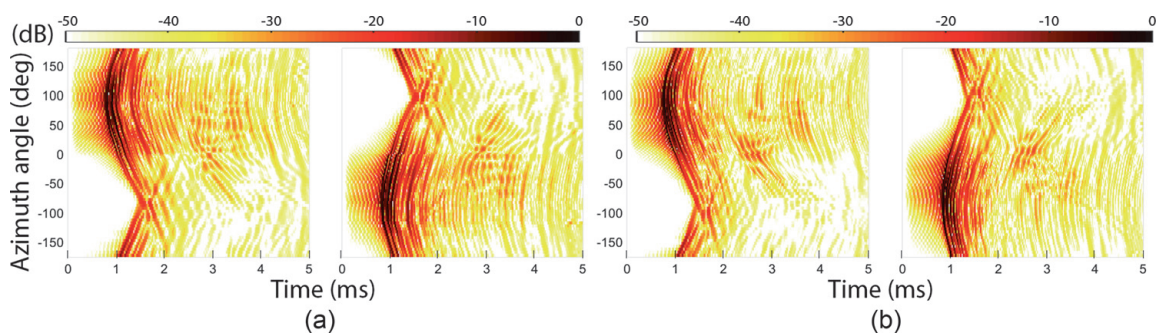
Humans can hear frequencies roughly between 20 Hz and 20 kHz, with frequencies at the lower end being perceived as vibrations or creaks, and with the

upper end decreasing with age and duration of noise exposure [52]. From the psychoacoustic perspective, frequencies down to 90 Hz contribute to sound lateralisation, i.e., localisation on the interaural axis within the head [53], and up to 16 kHz to sound localisation, i.e., localisation outside the head [54], defining the smallest frequency range for the HRTF acquisition. **Figure 2** shows the amplitude spectra of a binaural HRTF pair of two listeners. For each listener, the left and right columns show HRTFs of the left and right ear, respectively. The top row shows the HRTFs along the median, i.e., for the lateral angle of zero, from the front, via up, to the back. The bottom row shows the HRTFs along the Frankfurt plane, i.e., the horizontal plane located at the eye level. **Figure 2** demonstrates that HRTFs vary across ears, frequency, sound-source positions and listeners. The bottom panels emphasise the difference between ipsilateral and contralateral ear, showing the dynamic range, especially for frequencies higher than 6 kHz.

Assuming the propagation medium is air and a sonic speed of 340 m/s, the human hearing frequency range translates to wavelengths approximately between 1.7 cm and 17 m, resulting in different body parts affecting HRTFs in different frequency regions. The reflections of the torso create spatial-frequency modulations in the range of up to 3 kHz [1]. This effect can be observed in the top row of **Figure 2**, in the form of elevation-dependent spectral modulations along the median plane [55, 56]. Another contribution comes from the head, which shadows frequencies above 1 kHz. This effect can be observed in both rows of **Figure 2**, with large changes in the spectra beginning at around 1 kHz [57]. A large contribution is that of the pinna: The resonances and reflections within the pinna geometry create spectral peaks and notches, respectively, in frequencies above 4 kHz [54]. This effect can be observed in the bottom row of **Figure 2**.

From the perceptual perspective, the quality of these HRTF spectral profiles is important in many processes involved in spatial hearing. For example, sound-localisation performance deteriorates when these spectral profiles are disturbed by means of introducing spectral ripples [58], reducing the number of frequency channels [59] or spectral smoothing [60]. From the acoustic perspective, these spectral profiles show modulation depths of up to 50 dB [11], defining the required dynamic range in the process of HRTF acquisition.

The temporal aspects of HRTF acquisition are shown in **Figure 3** as the head-related impulse responses (HRIRs), i.e., HRTFs in the time domain, of the same listeners as in **Figure 2**. There are a few things to consider. First, the minimum length of the measurement is bounded by the length of the HRIRs. Their amplitude decays within the first 5 ms, setting the requirement for the room impulse response during the measurements [61]. After the 5 ms, the HRIRs decay below 50 dB, setting the requirement on the broadband signal-to-noise ratio (SNR) of the measurements. Further, because of the human sensitivity to interaural disparities, HRTF acquisition



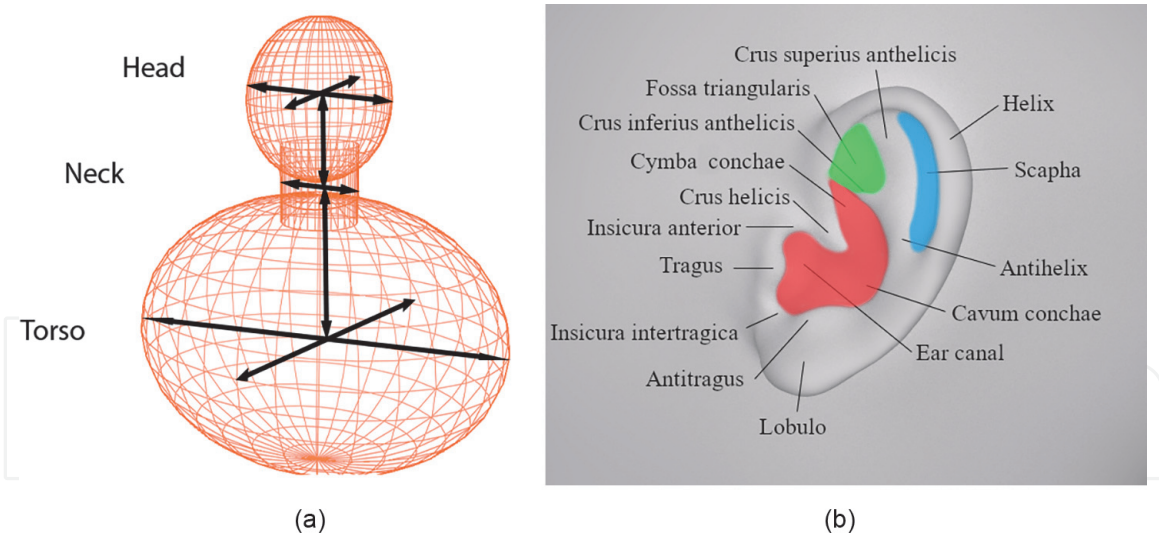
**Figure 3.**  
HRTF log-magnitudes in time domain along the eye-level horizontal plane for the same listeners as in **Figure 2**. Note the decay within the first 5 ms.

also requires an interaural temporal synchronisation. While sound sources placed in the median plane cause an ITD of zero (theoretically, reached only for identical path lengths to the two ears), just small deviations from the median plane cause potentially perceivable non-zero ITDs. Human listeners can detect ITDs being as small as  $10\ \mu\text{s}$  [53, 62], defining the interaural temporal precision required in the HRTF acquisition process. The ITD increases with the lateral angle of the sound source, reaching its extreme values for sources placed near the interaural axis [63, 64]. The largest ITD depends on the distance between the listener's two ears, mostly being defined by the listener's head width and depth [65], reaching ITDs of up to  $\pm 800\ \mu\text{s}$ . That ITD range translates to the sound's time of arrival (TOA) at an ear varying in the range of 1.6 ms, which needs to be considered in HRTF measurement by providing sufficient temporal space in the resulting impulse response.

HRTFs are continuous functions in space, even though, they are traditionally acquired for a finite set of spatial positions. From the *acoustic* perspective, assuming an HRTF bandwidth of 20 kHz, at least 2209 spatial directions are required to capture all spectro-spatial HRTF variations [66]. While this quite large number of spatial directions increases even further when considering multiple sound distances, it is in discrepancy with a smaller number of directions usually used in HRTF acquisition [11, 67–69]. One reason is the much smaller *perceptual* spatial resolution. From that perspective, the spatial resolution is limited by the ability to evaluate ITDs and changes in HRTF spectral profiles, both of which converge in the so-called minimum audible angles (MAAs). The MAA indicates the smallest detectable angle between two sound sources [70]. It depends on signal type [71, 72] and is minimal for broadband sounds [54, 73–75]. The MAA further depends on the direction of the source movement. Along the horizontal plane, the MAA can be as small as  $1^\circ$  for frontal sounds [76], increasing up to  $10^\circ$  for lateral sounds [77–79]. This translates to a high spatial-resolution requirement for frontal directions that can be relaxed with increasing lateral angle. Along the vertical planes, the MAA can be as low as  $4^\circ$  for frontal and rear sounds [76], increasing up to  $20^\circ$  for other sound directions [80]. Note that further relaxation of the requirement for spatial resolution can be achieved by using interpolation algorithms in the sound reproduction. For example, when using amplitude panning between the vertical directions [81], a resolution better than  $30^\circ$  does not seem to provide further advantages for localisation of sounds in the median plane [82]. Finally, when it comes to dynamic listening situations (involving listener or source movements), the MAAs further increase [83]. In order to account for sufficient spatial resolution when applying HRTFs in dynamic listening scenarios, the movement of the listener has to be monitored additionally to the modelling of sound source movement [84–86]. The minimum amount of directions and specific measurement points for a sufficiently sparse HRTF set are still current topics of research [87].

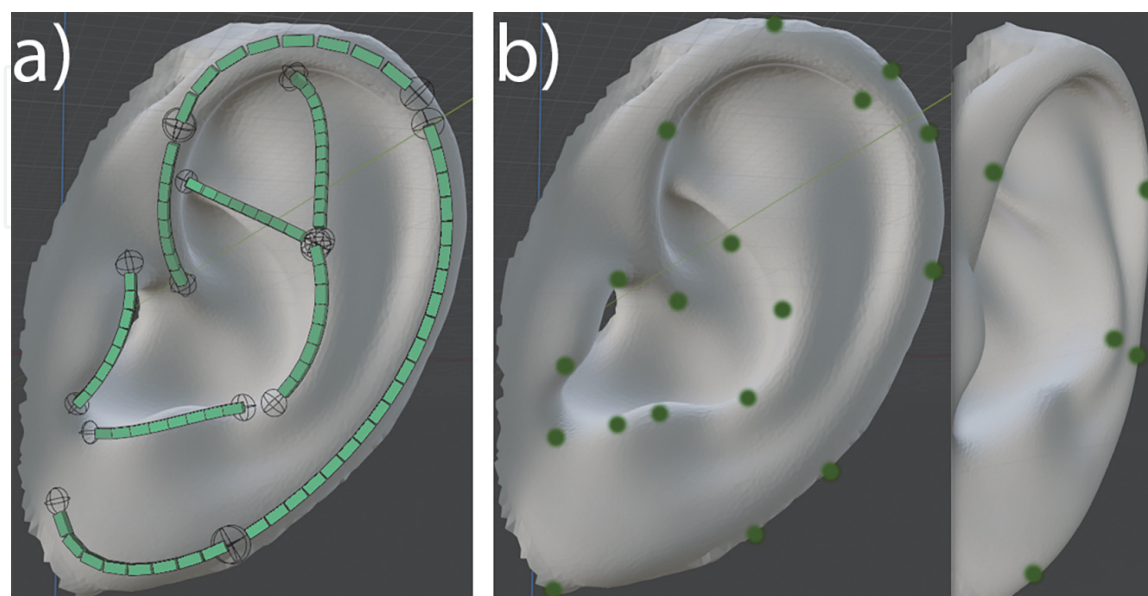
HRTFs are listener-specific, i.e., they vary among the listeners [21]. The reasons for that inter-individual variation are usually rooted in listener-specific morphology of the head and ears. For example, the variation in the head width of approximately  $\pm 2\ \text{cm}$  across the population causes variation in the largest ITD in the range of  $\pm 80\ \mu\text{s}$  [88]. **Figure 4** shows HRTF-relevant parts of the human body, where **Figure 4a** shows rough measures of the body and **Figure 4b** shows areas of the pinna responsible for the distinct spectral features in higher frequencies. The width and depth of head and torso have a large effect on HRTFs in the lower frequencies. The inter-individual variation in the pinnae geometry causes variations in HRTFs in frequencies above 4 kHz, with listener-specific differences of up to 20 dB [11]. The inter-individual variation in the HRTFs is rather complex because the pinna is a complex biological structure—small variations in geometry (in the range of millimetres) may cause drastic changes in HRTFs [90] along the vertical planes in high frequencies





**Figure 4.** HRTF-relevant parts of the human body. (a): Head and torso represented with simple shapes based on [57]. The black arrows denote the relevant measures. (b): Pinna and its distinctive regions. In red, green, and blue the concha, fossa triangularis, and scapha, respectively, denote the acoustically relevant areas [48, 56, 89].

[11], see **Figure 2**. However, not all pinna regions affect HRTFs equally [91]. Basically, the convex curvatures of the pinnae contribute to focusing the incoming sound waves towards the entry of the ear canals, comparable to a satellite dish. **Figure 4b** shows the anatomical areas important for localisation of sounds [48, 56, 89, 92, 93]. Currently, the description of the pinna geometry is not a trivial task. Pinnae have been described by means of anthropometric data stored in various data collections, e.g., [67, 69, 88, 94–96]. While the parameters used in these data collections do not seem to completely describe a pinna geometry from scratch, recent efforts aim at parametric pinna models able to generate non-pathological pinna geometries for arbitrary listeners [47, 48]. Such models describe the pinna geometry by means of various control points placed on the surface of a template pinna geometry. **Figure 5** shows two examples of the implementation of such



**Figure 5.** Examples of parametric pinna models. (a): Model from [47] consisting of Beziér curves (depicted in green), their control points (black spheres at both ends of a curve) and weights (not shown), linked to a template pinna geometry, (b): Model from [48], defined by control points of the pinna relief (green points) linked to proximal mesh vertices.



models. In **Figure 5a**, the pinna geometry is parametrised with the help of Beziér curves, i.e., polynomials within a spatial boundary [47]. **Figure 5b** shows a different approach; here, the parameterisation of the pinna is utilised with control points that move proximal local areas [48]. These parametric pinna models represent a step towards understanding the link between HRTFs and specific anatomical regions of the pinnae, and provide potential to synthesise large datasets of pinnae, e.g., in order to provide data for machine-learning algorithms.

In addition to the geometry, skin and hair may have an impact on HRTFs [97, 98] because of their direction-dependent absorption of the acoustic energy, especially at high frequencies. However, recent studies have shown that hair does not influence the localisation performance, but rather the perception of timbre instead [95, 99–101].

### 3. Acoustic measurement

The principle of an acoustic HRTF measurement relies on the system identification of the HRTF considered as a linear and time-invariant system. Here, an HRTF describes the propagation path between a microphone and a loudspeaker. Because of the binaural synchronisation, HRTFs are measured simultaneously at the two ears. The measurements are commonly performed for many source positions because of the required high spatial resolution. Recently, the details of the acoustic measurements, including a comprehensive list of HRTF measurement sites has been reviewed [20]. Thus, we only briefly introduce the basics and focus on the most recent advances in the acoustic HRTF measurement.

Typically, two omnidirectional microphones are placed in both ear canals, and the loudspeakers are arranged around the listener, ideally, with the number of loudspeakers corresponding to the number of HRTF positions to be measured. **Figure 6** shows two examples of measurement setups of various complexity: In **Figure 6a**, the listener is located on a turntable and moves within a fixed near-complete circular loudspeaker array. **Figure 6b** shows a similar approach with a near-complete spherical loudspeaker array, and **Figure 6c** shows the placement of a microphone in the ear canal so that it is membrane lines up with the entrance of the



**Figure 6.**

(a) Example of an HRTF measurement setup with mechanical rotation required. Listener sits on a chair (mounted on a turntable) surrounded by a loudspeaker arc (22 loudspeakers ranging from  $-30$  to  $210^\circ$  in  $5^\circ$ -steps). A head-tracker mounted on the head of the listener tracks head movements, triggering the need for measurement repetition in case of too large movements. (b) Example of more recent HRTF setups. 91 loudspeakers are mounted in a near-complete spherical array reducing the total measurement duration. (c) Example for a microphone placement in an HRTF measurement. Note the closed ear canal and the head-tracker sensor.

ear canal. Actually, it does not matter whether the microphones or loudspeakers are placed in the ear canal—this approach of ‘reciprocity’ is usually facilitated in numeric HRTF calculations (Section 4.4). However, setups with loudspeakers in the ears [102] lack signal-to-noise ratio (SNR) as the amplitude of the source signal needs to be low enough to not harm the listener, making the setup impractical for experiments. With the microphones in the ears, the most simple setups consist of a single loudspeaker moved around the listener [103]. Unfortunately, such setups lead to a long measurement duration for a dense set of HRTF positions. With the increasing availability of multichannel sound interfaces and adequate electroacoustic equipment, over the decades, the number of actually used loudspeakers increased. Setups with only a single loudspeaker moving around the listener have been replaced by setups with loudspeaker arcs surrounding the listener. In those setups, the listener sits on a turntable and either the listener (e.g., **Figure 6a**) or the loudspeaker arc is rotated [88, 104].

Recent approaches follow one of two different directions; On the one hand, generic and individual HRTFs are measured with a growing number of loudspeakers used in specialised facilities [67, 95]. Some even with such a large amount of loudspeakers that the listener is rotated for a few discrete positions, and post-processing algorithms interpolate between HRTF directions, e.g., the setup in **Figure 6b**. On the other hand, user-friendly individual HRTF measurement approaches are suggested, showing a trend towards decreasing the complexity of the measurement setup and using widely available equipment. In these approaches, only a single speaker is used and the listener is asked to move the head until a dense setup of HRTF directions can be obtained. These measurements enable simple systems to be used at home [105, 106], in which a head-tracking system records the listener’s head movements in real time and adapts the measured spatial HRTF grid. Head-above-torso orientations have to be considered additionally [100], but they reduce the complexity of the measurement setup and enable using widely available equipment, e.g., a commercially available VR headset and one arbitrary loudspeaker, in regular rooms, thus increasing the user-friendliness for setups [105].

Most of those recent approaches consider spatially discrete positions of the listener and/or the loudspeakers. In order to tackle the trade-off between high spatial resolution and long measurement duration, other recent advances have been made towards spatially continuous measurement approaches [107–109]. These approaches enable the measuring of all directions around the listener for a single elevation within less than 4 minutes [110]. Certainly, an advantage of such an approach is the access to the spatially continuous information, which is important especially for frontal HRTF directions. With more and more silent turntables and swivelled chairs, achieving a high SNR is not a big issue. Most recent approaches related to the spatially continuous measurement utilise Kalman filters to acquire system parameters representing HRTFs, and thus speed up the HRTF measurement in a multi-channel setup [111]. Compared to spatially discrete approaches, the spatially continuous method can achieve accuracy within a spectral error of 2 dB [109].

The requirements of the room are not rigorous: In principle, the measurement room does not have to be perfectly anechoic, but it has to fulfil some requirements regarding size and reverberation time. Room modes may exist below 500 Hz as they can be neglected in that frequency range [1]. Acceptable measurement results can be obtained as long as the first room reflection arises after 5 ms such that the measured room impulse responses can be truncated without truncating the HRIRs. Medium and large surfaces, i.e., the mount of the loudspeakers, the loudspeaker arc, the turntable, listener seat, etc., can potentially cause acoustical reflections overlapping with the direct sound path within the first 5 ms of the HRTF. These

reflections are usually damped, e.g., by covering the speakers in absorption material. Before the measurement, the listener's head has to be aligned in the measurement setup, adjusting the ears to the interaural axis of the system and the head to the Frankfurt plane. This alignment can be supported by, e.g., a laser system. The orientation and position of the listener's head should be monitored throughout the measurement procedure in order to detect listener's unwanted movements or position drifts. This helps when having to repeat potentially corrupted measurements.

The loudspeakers used for the measurements need to show a fast impulse response decay; fast enough to not interfere with the temporal characteristics of the HRTFs. This can be achieved by using loudspeaker drivers with light membranes, simple electric processing and no acoustic feedback such as a bass-reflex system. The acoustic short-circuit usually limits the lower frequency range of the loudspeakers, and multidriver systems are a common solution to that problem. In order to achieve a spatially compact acoustic source in a multidriver system, it is common to use coaxial loudspeaker drivers with an omnidirectional directivity pattern in HRTF measurement systems [112].

The placement of the microphones can also be an issue. Early setups used an open ear canal where the microphones were positioned close to the eardrum [11]. However, the effect of the ear canal does not seem to be direction-dependent, and its consideration in the measurement introduces technical difficulties and a large measurement variance [19, 113, 114]. Nowadays, the microphones are usually placed at the entrance of the ear canal which is acoustically blocked [11, 20]. Blocking the ear canal can be achieved by using microphones enclosed in earplugs made from foam or silicone or by wrapping the microphone in skin-friendly tape before inserting it. Note that such a measurement captures all directional-dependent features of the acoustic filtering by the outer ear, however, the directional-independent filtering by the ear canal is not captured. All cables from the microphone have to be flexible enough to minimise their effect on the acoustics within the pinna—one way is to lead the cable through the incisura intertragica and secure it with tape on the cheek, see **Figure 6c**.

In general, system identification can be performed with a variety of excitation signals. While previously Golay codes or other broadband signals have been used [115], more recently, the multiple exponential sweep method (MESM) [112] has been established and further improved [116], enabling fast HRTF measurement at high SNRs, reducing the discomfort for the listener. Still because of the imperfections in the electro-acoustic setup, a reference measurement is required to estimate the basis of the measurement without the effect of the listener, i.e., to estimate  $p_0$ . It is typically done for each microphone by placing the microphone in the centre of the measurement setup and recording the loudspeaker-microphone impulse response for all loudspeakers. The reference measurement can also be used to control the sound pressure level in order to avoid clipping at the microphones and analogue-digital converters. This can especially happen when each loudspeaker is driven within its linear range, but the overlapping signals from multiple loudspeakers raise the total level to ranges beyond the linear range of the recording system. Additionally, because of the HRTF's resonances, the level during the actual HRTF measurements can be up to 20 dB higher than that during the reference measurement, translating to a requirement for the headroom of at least 30 dB at the reference measurement. The maximum level is not only limited by the equipment; the listener's hearing range also needs to be considered, i.e., the maximum sound pressure level must neither create discomfort for the listener, nor go beyond the levels of safe listening. There is no special requirement for the listener regarding their audible threshold, hearing range or the visual sense. However, a particular

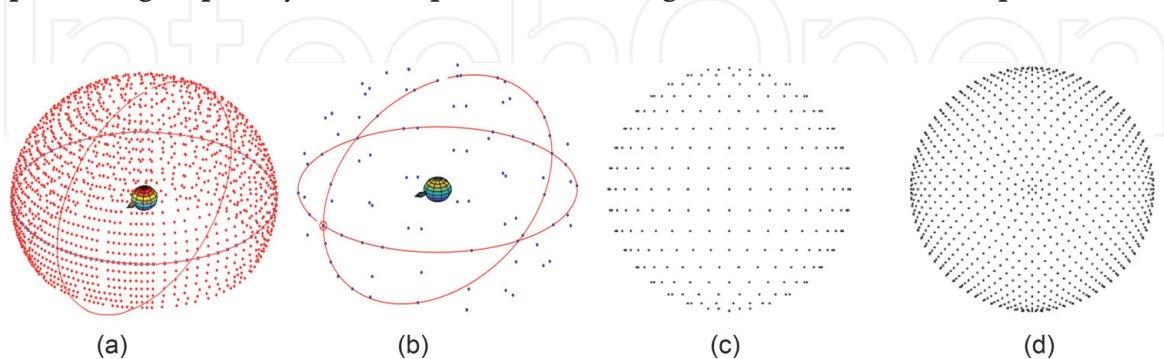


measurement equipment or a particular lab could have some restrictions on, e.g., the listener's weight or height.

**Figure 7** shows measurement grids of three exemplary setups and one measurement grid of a simulation setup. **Figure 7a** and **b** correspond to the measurement setups in **Figure 6a** and **b**. In these setups, not every loudspeaker plays a stimulus at every position around the listener. An extreme case is a loudspeaker positioned at  $90^\circ$  elevation that needs to be only measured once. **Figure 7c** shows another setup with uniformly distributed measurement points, and **Figure 7d** shows a uniform sampling grid used in numerical calculation experiments [95].

The repeatability of the measurement is an important issue. Within a single laboratory, changes in the room conditions such as temperature and humidity, as well as changes in the setup such as the ageing of the equipment may compromise the repeatability of the HRTF measurement [11, 20]. When comparing the HRTFs measurement across the labs, differences in the setups play also a role. In inter-laboratory and inter-method HRTF measurement comparison obtained for the same artificial head, severe ITD variations of up to  $200 \mu$  have been found [63, 64].

Once the HRTFs have been measured for all source positions, post-processing needs to be done before the HRTFs are ready to be used. First, in order to account for acoustic artefacts caused by the measurement room, a frequency-dependent windowing function is usually applied truncating the HRIRs [100, 117, 118]. Second, the measured HRIRs are equalised by the impulse response obtained from the reference measurements, i.e., with the microphone placed at the centre of the coordinate system with the listener absent. This equalisation can be either free-field or diffuse-field. For the free-field equalisation, the reference measurement is required only for the frontal direction ( $0^\circ$  azimuth,  $0^\circ$  elevation) [54], whereas for the diffuse-field equalisation, the reference measurement is the root mean square (RMS) impulse response of all directions [75], and the results are commonly denoted as directional transfer functions (DFT) [119]. Third, in most common rooms and even in (semi)anechoic rooms, reflections (or room modes) cause artefacts below 400 Hz, confounding the free-field property of HRTFs. Additionally, most loudspeakers used in the measurement are not able to reproduce low frequencies with sufficient power. Since the listener's anthropometry has a small effect on HRTFs in the low-frequency range, HRTFs can be extrapolated towards lower frequencies with a constant magnitude and linear phase [20, 117]. Further post-processing steps may include spectral smoothing to account for listener position



**Figure 7.** Four examples of spatial HRTF grid resolutions. (a) Almost spherical loudspeaker arc with moving listener, see also **Figure 6a**; the loudspeaker arc consists of 22 loudspeakers and yields 1550 directional measurements. Notice the higher resolution in the front of the listener as opposed to the lower resolution in the lateral regions. (b) Sparse measurement grid in a nearly spherical 91-loudspeaker array, see also **Figure 6b**, yielding 451 directional measurements with 5 different listener positions. (c) Measurement grid with 440 uniformly distributed points, measured with a loudspeaker arc with 37 loudspeakers from the HUTUBS database [95]. (d) Sampling grid for numerical calculations with 1730 directions. (c) and (d) are taken from the HUTUBS database [95].

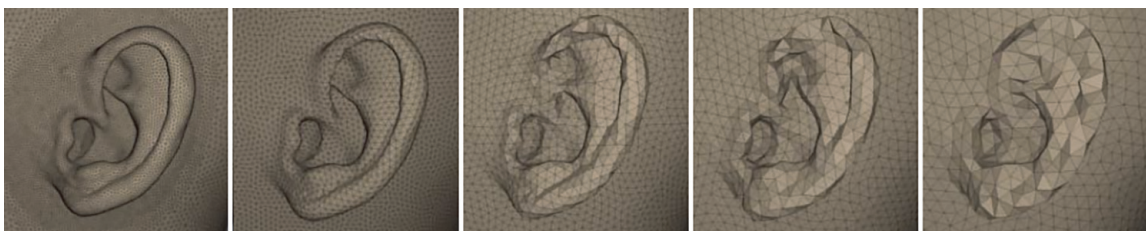
inaccuracies [60, 120] or adding a fractional delay to account for temperature changes followed by onset changes of the time signals [100].

The availability of acoustical HRTF measurements was a big step towards personalised binaural audio and virtual reality experience. However, even a fast or continuous measurement method requires the listener to sit still for a few minutes [104, 110, 112] in a specialised lab facility. Recent advances have been made towards both large-scale high-resolution and small-scale at-home easy-to-use solutions, providing HRTF acquisition to a large audience. Still, the imperfections in the electro-acoustic equipment set drawbacks of the acoustic measurement. Here, recent advances in the numeric calculations of the HRTFs can provide an interesting alternative.

#### 4. Numerical calculation of HRTFs

Generally, the calculation of HRTFs simulates the effects of the pinna, head and torso on the sound field at the eardrum. The goal is to numerically obtain the sound pressure at the two ears for a given set of frequencies and spatial positions. There are many methods to simulate wave propagation [121]. When applied to the HRTF calculation, all of the methods require a geometric representation of head and pinnae as input. For an accurate set of HRTFs, an exact 3D representation of the geometry, especially that of the pinnae with all their crests and folds, is of utmost importance [90]. The 3D geometry is represented using a discrete and finite set of elements, further denoted as ‘mesh’. A mesh is a representation of the region of interest (ROI), i.e., the object’s volume and surface, with the help of simple geometric elements. In most applications, the faces of these elements are assumed to be flat, which in turn explains the preference for triangular faces because they are always flat and therefore have one unique normal vector. This is not always the case for other shapes, e.g., quadrilaterals.

The requirements on the mesh have to consider geometrical as well as acoustical aspects. From the acoustic perspective, a typical rule of thumb for numerical calculation requires the average edge length (AEL) of elements to be at least a sixth of the smallest wavelength [122], which corresponds to an AEL of 3.5 mm for frequencies up to 16 kHz. However, in order to describe the pinna geometry sufficiently accurate, the average edge length (AEL) of the elements in the mesh needs to be around 1 mm, independently of the calculation method [90]. Some numerical calculation algorithms are, in general, more efficient and stable if the geometries are represented locally with elements of similar sizes and as regular as possible, e.g., almost equilateral triangles. To this end, the mesh may undergo a so-called *remeshing* [123], which inserts additional elements and resizes all elements to a similar size. **Figure 8** shows the same pinna in all panels, represented by meshes with increasing AELs from left to right.

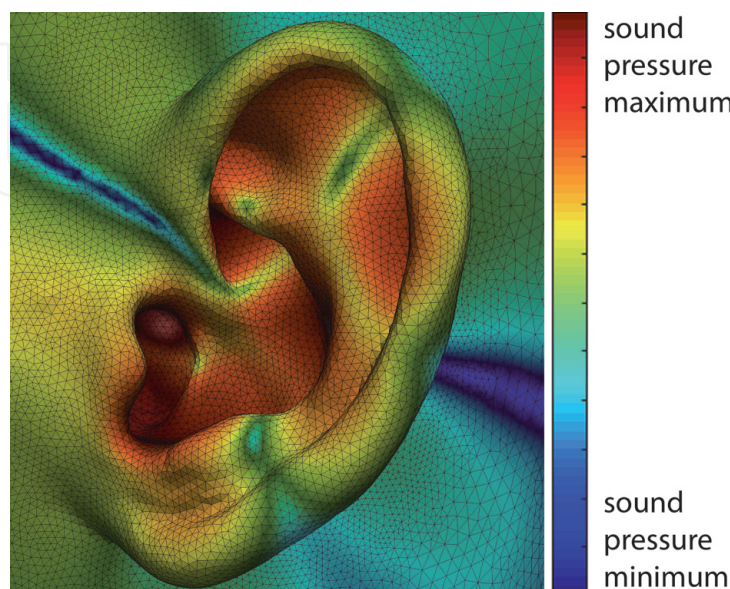


**Figure 8.**

*Pinna meshes represented by various AELs [90]. From left to right: AEL of 1 mm, 2 mm, 3 mm, 4 mm and 5 mm. Note how the representations of the helix and fossa triangularis degrade with increasing AEL.*

Interestingly, only the pinna regions contributing to the HRTF (compare **Figure 4b**) require to be accurately represented [56] and the remainder of the geometry can be more roughly modelled. This applies especially to the head, torso and neck, which can be represented by larger elements. These anatomical parts can additionally be approximated by simple geometric shapes, e.g., a sphere for the head, a cylinder for the neck and a rectangular cuboid or an ellipsoid representing the torso [65], see e.g., **Figure 4a**. To emphasise the sophisticated direction dependency of the pinna, **Figure 9** shows the calculated sound pressure distribution over the surface of the pinna. This simulation is calculated by defining one element in the centre of the ear canal as a sound source and evaluating the resulting sound pressure field at the vertices of the rest of the geometry; the procedure is explained thoroughly in Section 4.4.

The geometry can be captured via numerous approaches [124]: a laser scan [125], medical imaging techniques such as magnetic resonance imaging (MRI) [69, 126] and computer tomography (CT) [90], or photogrammetric reconstruction [127]. Laser, MRI and CT scans yield high-resolution meshes offering a small geometric error, but in turn, they need a special equipment. The laser scans are based on line-of-sight propagation and are able to measure short distances with an accuracy of up to 0.01 mm. The downside of line-of-sight propagation is that the manifolds of the pinnae are not easy to capture. In the medical imaging approaches, different downsides arise; acquiring the pinnae geometry via MRI is not a trivial process because they are flattened by the head support. This leads to two separate MRI measurements of each ear. The anatomy is then captured in ‘slices’ that can be stitched together in the postprocessing rather easily. The CT captures the anatomy in a similar way, but due to the high radiation exposure, such scans are usually not done with human subjects but with (silicone) mouldings of the listener’s ear. The overall procedure may take more time than an acoustic HRTF measurement and require the listener to either manufacture a moulding or meeting rather specific criteria for the scanning equipment (e.g., no tattoos, piercings, or implants). As an alternative, recent advances have been made for more widely applicable approaches such as photogrammetry [23, 128]. Photogrammetry is not only non-invasive but also can be done with widely available equipment, e.g., a smartphone or digital



**Figure 9.** Magnitude of the sound pressure calculated for each element of the surface for a 13-kHz sound source placed in the ear canal. Note the high dynamic range containing peaks (red) and notches (blue) in the distribution pattern in the area of the pinna.



camera, without having the listener to travel to a specialised facility. In a nutshell, the photogrammetrical approach works as follows: a set of photographs from different directions is made for each ear [127, 129], the so-called *structure from motion* [130] approach estimates the camera positions by analysing the mutual features across the photographs; a 3D point cloud is constructed; and a 3D mesh is created by connecting the points in the cloud. Note, that currently, manual corrections (e.g., smoothing to reduce noise, filling holes) are still required to reach the high quality of the meshes required for accurate HRTF calculations.

Simulations of acoustics require the information about the acoustic properties of the simulated objects. The HRTFs can be simulated with the 3D geometry represented as fully reflective, i.e., all surfaces having infinite acoustic impedance. With respect to localisation performance, only a small *perceptual* difference was found between acoustically measured and HRTFs calculated for acoustically reflective surfaces [101]. However, the impedance of various regions such as skin and hair may influence the direction-independent HRTF properties and cause changes in the perceived timbre [95, 99, 100].

In order to calculate HRTFs with sufficient spectral accuracy, the number of elements needs to be in the range of several tens of thousands, which might be important for the requirements of the computational power. Such large numerical problems usually require large amount of memory being in the range of Gigabytes. Nevertheless, the calculation time may reach a few days, especially when calculating HRTFs for many frequencies with high-resolution meshes. Note that if the used algorithm calculates HRTFs for each frequency independently, the calculations can be performed in parallel, and computer clusters can be used. This reduces the calculation time to a few hours for HRTFs the full hearing range and a mesh of several tens of thousands of elements.

All the algorithms for numerical HRTF calculation are based on the propagation of sound waves in the free field around a scattering object (also “scatterer”), usually described by the Helmholtz equation

$$\nabla^2 p(\mathbf{x}) + k^2 p(\mathbf{x}) = q(\mathbf{x}), \quad \mathbf{x} \in \Omega_e, \quad (2)$$

where  $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$  denotes the Laplace operator in 3D,  $p(\mathbf{x})$  denotes the (complex valued) sound pressure at a point  $\mathbf{x}$  for a given wavenumber  $k$  in the domain  $\Omega_e$  around the scattering object and  $q(\mathbf{x})$  denotes the (complex-valued) contribution of an external sound source at the acoustic field around the object. The wavenumber  $k$  is  $\frac{2\pi f}{c}$  with  $f$  being the frequency and  $c$  the speed of sound.

In order to solve the Helmholtz equation for a given scatterer, boundary conditions are necessary. The *Neumann* boundary condition assumes the object to be acoustically hard, and the (scaled) particle velocity at the boundary can be set to zero,

$$\frac{\partial p(\mathbf{x})}{\partial \mathbf{n}} = \nabla p(\mathbf{x}) \cdot \mathbf{n} = 0, \quad \mathbf{x} \in \Gamma,$$

where  $\mathbf{n}$  denotes the normal vector at the surface pointing away from the object. For the boundary condition at infinite distance, the so-called *Sommerfeld* radiation condition can be applied,

$$\left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \nabla p \right\rangle - ikp(\mathbf{x}) = o\left(\frac{1}{\|\mathbf{x}\|}\right), \quad \|\mathbf{x}\| \rightarrow \infty,$$

with  $o(\cdot)$  showing that the right side grows much faster than the left side. This ensures that the sound field decays away from the object [131].

For the calculation of HRTFs, the Helmholtz equation can be solved numerically by means of various approaches, which are based on a discretisation of the exterior domain  $\Omega_e$  around the scatterer  $\Gamma$ . Some of these methods solve the Helmholtz equation in the frequency domain, and others solve its counterpart, the wave equation, in the time domain. In general, the formulations and the results in the different domains can be transformed into each other by using, e.g., the Fourier transformation. In the following, we describe the most prominent methods used for HRTF calculations.

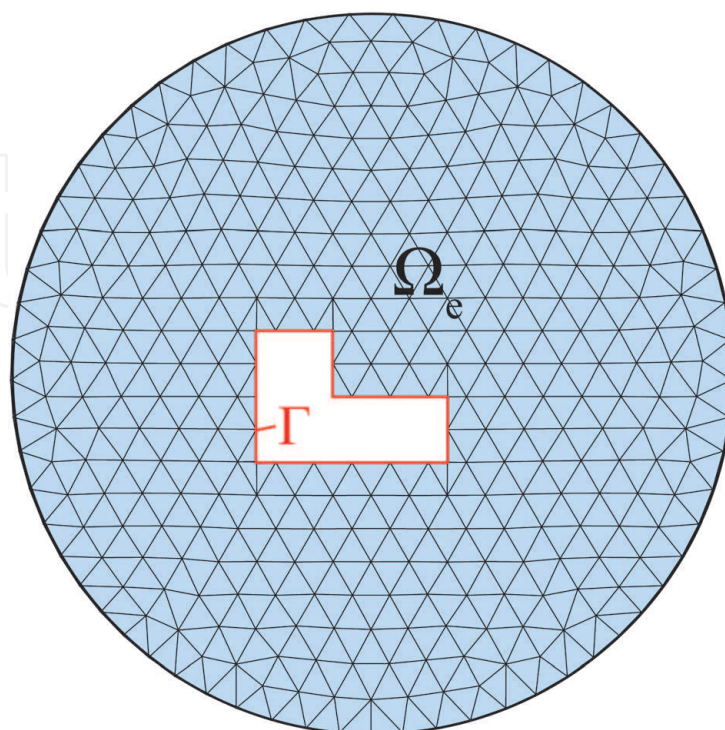
#### 4.1 The finite-element method

The finite-element method (FEM) solves the Helmholtz equation, Eq. (2), considering the scattering object or the domain around it as a volume [132]. **Figure 10** shows an example of a finite (domain) volume  $\Omega_e$  considered in the calculations with the scatterer with surface  $\Gamma$  placed inside that volume. To simulate the acoustic field around that object, the weak form of the Helmholtz equation is used, i.e., the equation is multiplied by a set of known test functions  $w(\mathbf{x})$  and integrated over the whole domain, thus transforming the partial differential equation (e.g., the Helmholtz equation) into an integral equation, that can be easier solved numerically:

$$\int_{\Omega_e} (\nabla^2 p(\mathbf{x}) + k^2 p(\mathbf{x})) w(\mathbf{x}) d\mathbf{x} = \int_{\Omega_e} q(\mathbf{x}) w(\mathbf{x}) d\mathbf{x}. \quad (3)$$

Secondly, the unknown pressure  $p(\mathbf{x})$  is approximated by a linear combination

$$p(\mathbf{x}) = \sum_{j=1}^N p_j \phi_j(\mathbf{x}) \quad (4)$$



**Figure 10.** 2D representation of meshes used in FEM. The elements are uniformly distributed and fitted to the boundary of the domain  $\Omega_e$ .

of so-called ansatz functions  $\phi_j(\mathbf{x}), j = 1, \dots, N$ . These ansatz functions, or element shape functions, help at interpolating between the discrete solutions for each point of the mesh. They are, in general, simple (real) polynomials defined locally on the elements of the mesh, e.g., having the value of 1 at their own coordinates and zero for other points of the mesh. Recent advances have been made towards adaptively finding higher-order polynomials and thus drastically reducing the computational effort [133, 134]. In theory, Eq. (3) should be fulfilled for all possible test functions  $w(\mathbf{x})$ , in practice, however, often the ansatz functions are also used as test functions, i.e.,  $w_i(\mathbf{x}) = \phi_i(\mathbf{x})$ . With this choice, Eq. (3) can be transformed into a linear system of equations

$$\mathbf{K}\mathbf{p} = \mathbf{g}, \quad (5)$$

where

$$\mathbf{K}_{ij} = \int_{\Omega_e} \frac{d\phi_i}{d\mathbf{x}} \frac{d\phi_j}{d\mathbf{x}} - k^2 \phi_i \phi_j d\mathbf{x}, \quad \mathbf{g}_i = \int_{\Omega_e} q(\mathbf{x}) \phi_i(\mathbf{x}) d\mathbf{x},$$

and  $\mathbf{p}$  is the vector containing the unknown coefficients  $p_i$  of the representation Eq. (4).

In general, the unknown coefficients  $p_i$  represent the complex sound pressure at a given node  $\mathbf{x}_i$  of the mesh. The integrals involved are solved using numerical methods [135].

When calculating HRTFs, the space around the scatterer is assumed to be continuous and infinite; in practice, this space has to be discretised and truncated to a finite domain by inserting a virtual boundary. When applied to the calculation of HRTFs, a virtual boundary of the (now finite) domain  $\Omega_e$  needs to be defined and conditions have to be set to avoid any reflections from this boundary, thus keeping in line with anechoic or free-field conditions. There are several methods to do so, with the so-called perfectly matched layers (PMLs) being the most popular among the reviewed HRTF calculation approaches. The PML is created by inserting an artificial boundary inside  $\Omega_e$ , e.g., a sphere with sufficiently large radius, and artificially damped equations are used to represent a solution that can then be numerically calculated, fulfilling the Sommerfeld radiation condition. Recent advances have been made to define PMLs automatically by extruding the boundary layer of the mesh and obtaining the geometric parameters during the extrusion step [136].

The FEM has been widely used in HRTF calculations [137–141] and yields similar results to acoustical HRTF measurements with spectral magnitude errors of approximately 1 dB [137, 141]. The downside, however, is the need to model 3D volumes around the head, resulting in models of a high number of elements, having a strong impact on the calculation duration.

## 4.2 The finite-difference time-domain method

A similar approach as the FEM can also be followed in the time domain. By using a short sound burst in the time domain as an input signal, the HRTFs within a wide frequency range can be calculated at once. This approach is called the finite-difference time-domain (FDTD) method [142] and can be derived by solving the wave equation in the time domain

$$\nabla^2 \tilde{p}(\mathbf{x}, t) - \frac{\partial^2 \tilde{p}(\mathbf{x}, t)}{\partial t^2} = \tilde{q}(\mathbf{x}, t),$$



where  $\tilde{p}$  and  $\tilde{q}$  denote sound pressure fields in the time domain. The PML is applied to create the boundary conditions for outgoing sound waves. The evaluation grid is sampled evenly in cells across the domain with grid spacing  $h$ , considering discrete time steps  $m$ . A key parameter for numerical stability of the FDTD is the Courant number

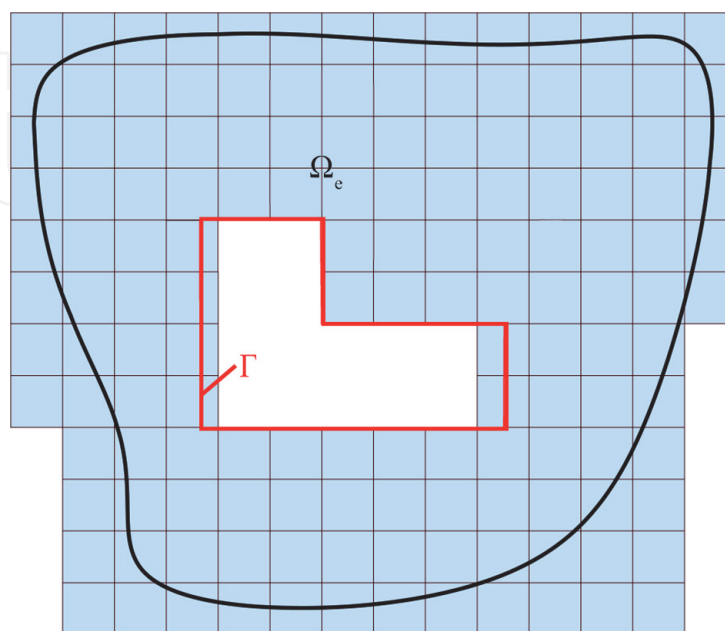
$$\beta = \frac{cm}{h},$$

defining the number of cells the sound propagates per time step. Typically, in order to obtain stable HRTF calculations, the Courant number is  $\beta = 1/\sqrt{3}$ .

**Figure 11** shows a 2D representation of a mesh used in the FDTD method. Note that because the mesh needs to consist of evenly spaced elements, most of the objects cannot be represented accurately and a sampling error is introduced at the boundary surface  $\Gamma$  of the object. Additionally, as derivatives of functions are approximated by finite differences, the arithmetic operations are valid for infinite resolution, but when calculating on physical computers, the precision depends on the number format used and the gridsize, introducing errors in the results [143]. Refining the grid is a potential solution to the sampling error for staircase approximations [144, 145], and when framing this problem to HRTF calculations, spectral magnitude errors of 1 dB up to 8 kHz and 2 dB up to 18 kHz can be achieved [146, 147], suggesting only negligible increase in localisation errors when listening to HRTFs calculated by the FDTD method.

Because of the additional sampling errors for irregular domains, recent advances have been made towards using quasi-cartesian grids [148], dynamically choosing grid resolutions [149], or towards the finite-volume method (FVTD), which is based on energy conservation and dissipation of the system as a whole and uses the integral formulation of the FDTD [150]. One solution approach there is to adaptively sample the grid at the boundary and introduce unstructured or fitted cells [151, 152]. A thorough comparison between FEM, FDTD and FVTD methods is available in [153].

In fact, the FDTD method has been widely applied to HRTF calculations [145, 146, 154, 155], and it certainly offers the advantage of calculating broadband



**Figure 11.** 2D representation of meshes used in the FDTD method. Note that in this representation, the object surface  $\Gamma$  does not line up exactly with the sampling grid.

HRTFs while not introducing additional computational cost when multiple inputs or outputs are used. However, because of the complex geometry of the pinnae, a submillimetre sampling grid is required, resulting in the need for a delicate preprocessing.

### 4.3 The boundary-element method

The boundary element method (BEM) is based on a special set of test functions in the weak formulation of the Helmholtz equation Eq. (3), namely the Green's function

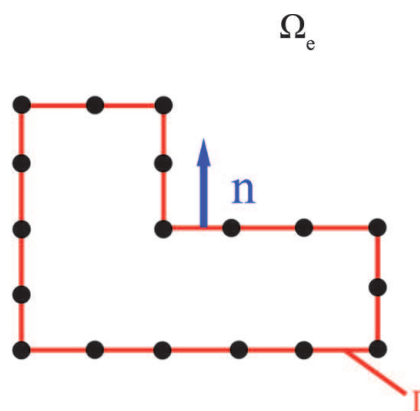
$$G(\mathbf{x}, \mathbf{y}) = \frac{e^{ikr}}{4\pi r},$$

where  $r = \|\mathbf{x} - \mathbf{y}\|$ , and  $\mathbf{x}$  and  $\mathbf{y}$  are two points in space. By using this function, it is possible to reduce the weak form of the Helmholtz equation to an integral equation, i.e., the boundary integral equation (BIE), that only involves integrals over the surface  $\Gamma$  of the object, and *not* the volume  $\Omega_e$ . Assuming that the external sound source as a point source at  $\mathbf{x}^*$ , and an acoustically reflecting (= sound hard,  $\nabla p(\mathbf{y}) \cdot \mathbf{n} = 0, \mathbf{y} \in \Gamma$ ) surface, the sound field at a point  $\mathbf{x}$  on a smooth part of that surface  $\Gamma$  is given by:

$$\frac{1}{2}p(\mathbf{x}) - \int_{\Gamma} H(\mathbf{x}, \mathbf{y})p(\mathbf{y})d\mathbf{y} = G(\mathbf{x}^*, \mathbf{x})p_0, \quad (6)$$

where  $H(\mathbf{x}, \mathbf{y}) = \frac{\partial G}{\partial \mathbf{n}_y}(\mathbf{x}, \mathbf{y})$  is obtained by the derivative of the Green's function at a point  $\mathbf{y}$  in the direction of vector  $\mathbf{n}$  normal to the surface at this position, and  $p_0$  denotes the strength of the sound source.

In comparison with the other two methods, the BEM has the advantage that only the *surface* of the object such as the head and the pinnae needs to be discretised, whereas in FEM and the FDTD method also a discretisation of the *volume* surrounding the head has to be considered, see **Figures 10–12**. Thus, in the boundary element method, all calculations can be reduced to a manifold described in 2D, in our case, the domain of interest is reduced to the surface of the head. A second advantage of the BEM is that by using the Green's function, the Sommerfeld radiation condition is automatically fulfilled. Additionally, no domain boundary has to be



**Figure 12.** 2D representation of a BEM mesh. Note that only the boundary of the surface  $\Gamma$  is considered and the domain volume  $\Omega_e$  does not have to be sampled.

introduced, such as the PML. This renders the BEM an attractive method for calculating sound propagation in infinite domains, i.e., in free-field, as is the assumption when calculating HRTFs [156].

In order to solve a BEM problem, the BIE is discretized and solved by using methods such as the Galerkin, collocation or Nyström [157–159], all with the common goal of yielding a linear system of equations.

For the Galerkin method, the unknown pressure is approximated by a linear combination of ansatz functions as in Eq. (4). The BIE is again multiplied with a set of test functions (similar to the test functions  $\phi_i(\mathbf{x})$  used in FEM) and integrated with respect to  $\mathbf{x}$  yielding a linear system of equations as in Eq. (5), where

$$\mathbf{K}_{ij} = \int_{\Gamma} \frac{\phi_i(\mathbf{x})\phi_j(\mathbf{x})}{2} d\mathbf{x} - \iint_{\Gamma} \phi_i(\mathbf{x})H(\mathbf{x}, \mathbf{y})\phi_j(\mathbf{y})d\mathbf{y}d\mathbf{x},$$

and

$$\mathbf{g}_i = \int_{\Gamma} G(\mathbf{x}_0 - \mathbf{x})\phi_i(\mathbf{x})d\mathbf{x}.$$

Another commonly used approach especially used in engineering is collocation with constant elements, i.e., the sound field is assumed to be constant on each element of the mesh, and the BIE is solved at the midpoints  $\mathbf{x}_i$  of each element (the set of all  $\mathbf{x}_i$  are called collocation nodes) yielding a linear system of equations as in Eq. (5), where

$$\mathbf{K}_{ij} = \frac{1}{2} \delta_{ij} - \int_{\Gamma_j} H(\mathbf{x}_i, \mathbf{y})d\Gamma, \quad \mathbf{g}_i = G(\mathbf{x}^*, \mathbf{x}_i)p_0.$$

$\mathbf{p}_i = p(\mathbf{x}_i)$  and with  $\mathbf{x}^*$  being the position of the sound source outside the head. The integrals over each element are numerically solved using appropriate quadrature formulas (weighted sum of function values) and

$$\delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$$

The BIE is solved for a given set of frequencies and the solutions  $\mathbf{p}_i$  at the collocation nodes are then used to derive the HRTFs. Note that the collocation method can be interpreted as the Galerkin method utilising the delta functionals  $\delta(\mathbf{x}_i - \mathbf{x})$  as test functions. A thorough comparison between Galerkin and collocation approaches can be found in [160].

The discretisation of just the surface introduces additional challenges. First, the Green's function becomes singular at the boundary where  $\|\mathbf{x} - \mathbf{y}\| = 0$  and special quadrature formulas need to be used close to these singularities [161, 162]; and second, the system matrix  $\mathbf{K}$ , although small, is usually densely populated, which poses a challenge for computer memory and the efficiency of the linear solver used. When considering HRTF calculations for frequencies up to 20 kHz, high-resolution meshes are required and the corresponding linear systems may contain up to 100,000 unknowns.

In order to efficiently deal with such large systems, the BEM can be coupled with methods speeding up matrix–vector multiplications, such as the fast-multipole method (FMM) [163] or  $\mathcal{H}$ -matrices [164] (so-called ‘hierarchical’ matrices). These methods have enabled an efficient and feasible calculation of HRTFs [165]. In a



nutshell, these methods aim at providing a method for an efficient and fast matrix–vector multiplication and are based on two steps. First, the elements of the mesh are grouped into clusters of approximately the same size with cluster midpoints  $\mathbf{z}_i$ . Second, for two clusters  $\mathcal{C}_1$  and  $\mathcal{C}_2$ , that are sufficiently apart from each other, a separable approximation of the Green’s function

$$G(\mathbf{x}, \mathbf{y}) \approx G_1(\mathbf{x} - \mathbf{z}_1)M(\mathbf{z}_1, \mathbf{z}_2)G_2(\mathbf{y} - \mathbf{z}_2)$$

is found. This approximation has two advantages: the local expansions  $G_1$  and  $G_2$  have to be made only once for each cluster, and the interaction between the elements of different clusters is reduced to a single interaction of the cluster midpoints. The resulting linear system of equations is then solved using an iterative equation solver [166].

Although the Helmholtz equation for external problems has a unique solution at all frequencies, the BIE has uniqueness problems at certain critical frequencies [159, 167]. Thus, to avoid numerical problems, the BEM needs to be stabilised at these frequencies, e.g., by using the Burton-Miller method [167]. BEM has been widely used to calculate HRTFs [165, 168–171] analysing the process from various perspectives. When applied to an accurate and high-resolution representation of the pinna geometry, BEM can yield similar results to the acoustic HRTF measurements by means of sound localisation performance [101, 172].

#### 4.4 Reciprocity

In principle, in order to calculate an HRTF set, the Helmholtz equation needs to be solved for every source position  $\mathbf{x}_j^*$  separately, leading in up to several thousand right-hand sides in Eq. (5). Solving that many equations cannot be done quickly even with the help of iterative solvers. On the other hand, the HRTF calculation for the second ear is quite simple because the solution obtained from the solver is already available for every element of the surface, i.e., at the element representing the ear canal of the second ear. The approach of reciprocity can help to significantly speed up the calculations by swapping the role of the many source positions with the two elements representing the ear canals, requiring us to solve Eq. (5) only twice, i.e., once for each of the ears.

Helmholtz’ reciprocity theorem states that switching source and receiver positions do not affect the observed sound pressure. When applied to HRTF calculations, virtual loudspeakers are placed in the entrance of the ear canal (replacing the virtual microphones) and the many simulated sound sources are represented by many virtual microphones (replacing the many virtual loudspeakers around the listener). By doing so, the computationally expensive part of the BEM, i.e., solving a linear system of equations to calculate the sound pressure at the surface, needs to be done only twice, namely once for each ear. Subsequently, the sound pressure at positions around the head can be calculated fairly easy and efficiently.

In more detail, assume that a point source with strength  $p_0$  at the position  $\mathbf{x}_j^*$  causes a mean sound pressure of  $\bar{p}$  over a small domain with area  $A_0$  at the entrance of the ear canal. If the domain is sufficiently small, the mean sound pressure is an accurate representation of the actual sound pressure in this domain. By applying the reciprocity, we introduce a reciprocal sound source at the entrance of the ear canal which introduces a velocity  $v_0$  and then calculate the sound pressure  $p(\mathbf{x}_j^*)$  at the actual sound-source position around the listener. The pressures  $p(\mathbf{x}_j^*)$  and  $\bar{p}$  are linked by

$$\bar{p} = \frac{p_0 p(\mathbf{x}_j^*)}{A_0 v_0}.$$

The reciprocal sound source can be modelled by vibrating elements  $\Gamma_{\text{vib}}$ , i.e., elements with an additional velocity boundary condition

$$v(\mathbf{y}) = \frac{1}{i\omega\rho} \frac{\partial p}{\partial \mathbf{n}} = v_0, \quad \mathbf{y} \in \Gamma_{\text{vib}}$$

where  $\omega = 2\pi f$ , and  $\rho$  is the density of air. Note that  $v_0$  can be an arbitrary positive number because when calculating HRTFs [see for example Eq. (1)], the pressure is normalised by the reference pressure  $p_0$ , thus cancelling  $v_0$ . With this additional boundary condition, first, BEM can be used to calculate the sound field at the surface  $\Gamma$ , and then, Green's representation is applied to calculate the pressure at all positions of the actual sound sources  $\mathbf{x}_j^*$ ,

$$p(\mathbf{x}_j^*) - \int_{\Gamma} H(\mathbf{x}_j^*, \mathbf{y}) p(\mathbf{y}) d\mathbf{y} + i\omega\rho \int_{\Gamma} G(\mathbf{x}_j^*, \mathbf{y}) v(\mathbf{y}) d\mathbf{y} = 0.$$

Note that this equation is calculated after a discretisation, and because  $p(\mathbf{y})$  at the surface is known from the BEM solution, the calculation of the sound pressure around the head  $p(\mathbf{x}_j^*)$  is a simple matrix multiplication.

Reciprocity, combined with FMM-coupled BEM has been applied to calculate HRTFs, enabling calculations for a large spatial HRTF set within a few hours even on a standard desktop computer [172].

## 5. Other issues related to HRTF acquisition

Over decades, HRTFs have been collected and stored in databases. Such databases are important for educational aspects, training of neural network algorithms [34, 37] and further research [23, 25–28, 173]. While in the early HRTF research days, HRTFs have been stored by each lab in a different format, since 2015, the spatially oriented format for acoustics (SOFA) is available to store HRTFs in a flexible but well-described way facilitating an easy exchange between the labs and applications. SOFA is a standard of the Audio Engineering Society under the name AES69. SOFA provides a uniform description of spatially oriented acoustic data such as HRTFs, spatial room impulse responses, and directivities [15].

When it comes to anthropometric data, unfortunately, there is currently no common format to specify and exchange anthropometric data. This is partially because currently, it is not known, which data are important. Some laboratories use the CIPIC parameters [88], some have extended them [174], and others have created whole new sets of parameters [128, 175]. An overview of currently used anthropometric parameters can be found in [176]. The development of parametric pinna models may shed light on the relevance of parameters needed to be stored in the future. The listener's geometry can also be stored in non-parametric representations such as meshes and point clouds of listener's ears and head. To this end, typical 3D dataset formats are used, e.g., OBJ, PLY or STL. These formats are widely used in computer graphics and thus easily accessible by many corresponding applications. A large collection of HRTF databases stored in SOFA, with some of them

combined with meshes stored in OBJ, PLY and STL files is available at the SOFA website.<sup>1</sup>

When HRTFs are obtained, there is strong demand to evaluate their quality. This is especially interesting when comparing the results from numerical HRTF calculations. The evaluations can be performed at various levels: geometrical, acoustical and perceptive. The evaluation at the geometric level can be done by comparing the deviation between two meshes of the pinna and representing the deviation as the Hausdorff distance [177]. The evaluation at the acoustic level can be done by calculating the spectral distortion

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( 20 \log \frac{\widehat{HRTF}(\mathbf{x}^*, f_i)}{HRTF(\mathbf{x}^*, f_i)} \right)^2}, \quad (7)$$

where  $\widehat{HRTF}(\mathbf{x}^*, f_i)$  denotes the calculated and  $HRTF(\mathbf{x}^*, f_i)$  the measured one, summarised over  $N$  discrete frequencies. The evaluation at the perceptual level can be simulated by means of auditory modelling [50] or direct performance of localisation experiment [50, 90, 147]. Especially the evaluation of localisation errors in the median plane can be relevant because the sound localisation in the median plane is directly related to the quality of monaural spectral features in the HRTFs [46, 178]. A calculated HRTF set yielding similar perceptual results as the natural listener's HRTFs can be described as *perceptually valid*.

## 6. Conclusions

With a specialised measurement setup, acoustic HRTF measurements can be done within a few minutes. Still, such setups are expensive and require the listener to sit or stand still for the whole measurement duration. The requirement of specialised components has been limiting the popularity of the acoustic methods. Recent advances, however, have been made by integrating head-movement tracking in systems to be used at home, especially since the commercialisation of VR headsets. These advances provide an easy-to-use measurement setup, but still need investigation on how many and which measurement positions are crucial to acquire a sufficient measurement grid for perceptually valid HRTFs.

With the availability of numerical HRTF calculations, the acquisition of personalised HRTFs has undergone significant advances. While the acoustic HRTF measurement still remains the reference acquisition method, numerical HRTF calculation paves the road towards personalised HRTFs available for a wide audience. The most widely used approaches, FEM, FDTD, BEM and BEM coupled with the FMM, when applied under optimal conditions, can yield acoustically and perceptually valid results.

Machine learning and neural networks gain increasing popularity and, in the future, may even further push the usability of numerical HRTF calculations. For example, neural networks might be able to support the photogrammetric mesh acquisition or even estimate the HRTFs directly from listener-specific anthropometric data such as photographs. Further improvements in terms of efficiency, accuracy and precision are still ongoing subject of research.

Despite the clear definition when it comes to storing an HRTF data set by means of SOFA, a similar definition for the description of anthropometric data is still not

<sup>1</sup> <https://www.sofaconventions.org/mediawiki/index.php/Files>



available. This might be rooted in our poor understanding of the importance of parts of the pinna and its contribution to the HRTF. Here, a clear goal is to better understand the anthropometry and its relation with HRTFs. All this future work heads into the direction of expanding the access to personalised HRTFs enabling their availability for everyone.

## **Acknowledgements**

This work was supported by the Austrian Research Promotion Agency (FFG, project 'softpinna' 871263) and the European Union (EU, project 'SONICOM' 101017743, RIA action of Horizon 2020). We thank Harald Ziegelwanger for visualising the sound pressure in Figure 9.

## **Conflict of interest**


The authors declare no conflict of interest.

## **Author details**

Katharina Pollack\*, Wolfgang Kreuzer and Piotr Majdak  
Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria

\*Address all correspondence to: [katharina.pollack@oeaw.ac.at](mailto:katharina.pollack@oeaw.ac.at)

## **IntechOpen**

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Algazi VR, Avendano C, Duda RO. Elevation localization and head-related transfer function analysis at low frequencies. *The Journal of the Acoustical Society of America*. 2001; **109**(3):1110-1122. DOI: 10.1121/1.1349185
- [2] Batteau DW. The role of the pinna in human localization. *Proceedings of the Royal Society of London Series B. Biological Sciences*. 1967;**168**(1011):158-180. DOI: 10.1098/rspb.1967.0058
- [3] Baumgartner R, Reed DK, Tóth B, Best V, Majdak P, Colburn HS, et al. Asymmetries in behavioral and neural responses to spectral cues demonstrate the generality of auditory looming bias. *Proceedings of the National Academy of Sciences*. 2017;**114**(36):9743-9748, ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1703247114
- [4] Fisher HG, Freedman SJ. The role of the pinna in auditory localization. *Journal of Auditory Research*. 1968; **168**(1011):158-180
- [5] Hebrank J, Wright D. Spectral cues used in the localization of sound sources on the median plane. *The Journal of the Acoustical Society of America*. 1974; **56**(6):1829-1834. DOI: 10.1121/1.1903520
- [6] Musicant AD, Butler RA. The influence of pinnae-based spectral cues on sound localization. *The Journal of the Acoustical Society of America*. 1984; **75**(4):1195-1200. DOI: 10.1121/1.390770
- [7] Majdak P, Baumgartner R, Jenny C. Formation of three-dimensional auditory space. In: Blauert J, Braasch J, editors. *The Technology of Binaural Understanding, Modern Acoustics and Signal Processing*. Cham, ISBN: 978-3-030-00386-9: Springer International Publishing; 2020. pp. 115-149. DOI: 10.1007/978-3-030-00386-9\_5
- [8] Majdak P, Baumgartner R, Laback B. Acoustic and non-acoustic factors in modeling listener-specific performance of sagittal-plane sound localization. *Frontiers in Psychology*. 2014;**5**:319. DOI: 10.3389/fpsyg.2014.00319
- [9] Seeber BU, Fastl H. Subjective selection of non-individual head-related transfer functions. In: *Proceedings of the International Conference on Auditory Display*. Atlanta, Georgia: Georgia Institute of Technology; 2003. pp. 259-262
- [10] Wenzel EM, Arruda M, Kistler DJ, Wightman FL. Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America*. 1993;**94**(1):111-123. DOI: 10.1121/1.407089
- [11] Møller H, Sørensen MF, Hammershøi D, Jensen CB. Head-related transfer functions of human subjects. *Journal of the Audio Engineering Society*. 1995;**43**:300-321
- [12] Macpherson EA, Middlebrooks JC. Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *The Journal of the Acoustical Society of America*. 2002;**111** (5 Pt 1):2219-2236. DOI: 10.1121/1.1471898
- [13] Reijniers J, Vanderelst D, Jin C, Carlile S, Peremans H. An ideal-observer model of human sound localization. *Biological Cybernetics*. 2014;**108**(2):169-181, ISSN: 0340-1200. DOI: 10.1007/s00422-014-0588-4
- [14] Majdak P, Goupell MJ, Laback B. 3-d localization of virtual sound sources: Effects of visual environment, pointing method, and training. *Attention, Perception, & Psychophysics*. 2010; **72**(2):454-469. DOI: 10.3758/APP.72.2.454

- [15] Majdak P, Carpentier T, Nicol R, Roginska A, Suzuki Y, Watanabe K, et al. Spatially oriented format for acoustics: A data exchange format representing head-related transfer functions. In: Proceedings of the 134th Convention of the Audio Engineering Society (AES), Page Convention Paper 8880. Roma, Italy: Audio Engineering Society; 2013
- [16] Majdak P, Hollomey C, Baumgartner R. The auditory modeling toolbox. In: The Technology of Binaural Listening. Berlin, Heidelberg: Springer; 2021. pp. 33-56
- [17] Søndergaard P, Majdak P. The auditory modeling toolbox. In: Blauert J, editor. The Technology of Binaural Listening. Berlin-Heidelberg, Germany: Springer; 2013. pp. 33-56. DOI: 10.1007/978-3-642-37762-4\_2
- [18] Guezenoc C, Segulier R. HRTF individualization: A survey. In Audio Engineering Society convention 145, page Convention Paper 10129. New York, New York, United States: Audio Engineering Society; 2018
- [19] Hammershøi D, Møller H. Sound transmission to and within the human ear canal. The Journal of the Acoustical Society of America. 1996;**100**(1): 408-427. DOI: 10.1121/1.415856
- [20] Li S, Peissig J. Measurement of head-related transfer functions: A review. Applied Sciences. 2020;**10**(14): 5014. DOI: 10.3390/app101450140 Number: 14 Publisher: Multidisciplinary Digital Publishing Institute
- [21] Middlebrooks JC. Individual differences in external-ear transfer functions reduced by scaling in frequency. The Journal of the Acoustical Society of America. 1999;**106**(3): 1480-1492. DOI: 10.1121/1.427176
- [22] Iida K, Aizaki T, Kikuchi T. Toolkit for individualization of head-related transfer functions using parametric notch-peak model. Applied Acoustics. 2022;**189**:108610. DOI: 10.1016/j.apacoust.2021.108610
- [23] Torres-Gallegos EA, Orduna-Bustamante F, Arámbula-Cosío F. Personalization of head-related transfer functions (HRTF) based on automatic photo-anthropometry and inference from a database. Applied Acoustics. 2015;**97**:84-95. DOI: 10.1016/j.apacoust.2015.04.009
- [24] Guezenoc C, Segulier R. A wide dataset of ear shapes and pinna-related transfer functions generated by random ear drawings. The Journal of the Acoustical Society of America. 2020; **147**(6):4087-4096. DOI: 10.1121/10.0001461
- [25] Jin CT, Zolfaghari R, Long X, Sebastian A, Hossain S, Glaunés J, et al. Considerations regarding individualization of head-related transfer functions. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary, AB, Canada: IEEE; 2018. pp. 6787-6791. DOI: 10.1109/ICASSP.2018.8462613
- [26] Lu D, Zeng X, Guo X, Wang H. Personalization of head-related transfer function based on sparse principle component analysis and sparse representation of 3d anthropometric parameters. Australia: Acoustics; 2019. pp. 1-10. DOI: 10.1007/s40857-019-00169-y
- [27] Tommasini FC, Ramos OA, Hüg MX, Bermejo F. Usage of spectral distortion for objective evaluation of personalized hrtf in the median plane. International Journal of Acoustics & Vibration. 2015;**20**(2):81-89
- [28] Zhang M, Ge Z, Liu T, Wu X, Qu T. Modeling of individual HRTFs based on spatial principal component analysis. IEEE/ACM Transactions on Audio,



- Speech, and Language Processing. 2020; 28:785-797. DOI: 10.1109/TASLP.2020.2967539
- [29] Zhang M, Kennedy R, Abhayapala T, Zhang W. Statistical method to identify key anthropometric parameters in HRTF individualization. In: 2011 Joint Workshop on Hands-free Speech Communication and Microphone Arrays. Edinburgh, Scotland: IEEE; 2011. pp. 213-218. DOI: 10.1109/HSCMA.2011.5942401
- [30] Hu H, Zhou L, Zhang J, Ma H, Wu Z. Head related transfer function personalization based on multiple regression analysis. In: 2006 International Conference on Computational Intelligence and Security. Vol. 2. Guangzhou, China: IEEE; 2006. pp. 1829-1832. DOI: 10.1109/ICCIAS.2006.295380
- [31] Huang Q, Zhuang Q. HRIR personalisation using support vector regression in independent feature space. *Electronics Letters*. 2009;45(19): 1002-1003
- [32] Zolfaghari R, Epain N, Jin CT, Glaunes J, Tew A. Large deformation diffeomorphic metric mapping and fast-multipole boundary element method provide new insights for binaural acoustics. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). London: IEEE; 2014. pp. 2863-2867. DOI: 10.1109/ICASSP.2014.6854123
- [33] Grijalva F, Martini LC, Florencio D, Goldenstein S. Interpolation of head-related transfer functions using manifold learning. *IEEE Signal Processing Letters*. 2017;24(2):221-225. DOI: 10.1109/LSP.2017.2648794
- [34] Gebru ID, Marković D, Richard A, Krenn S, Butler GA, De la Torre F, et al. Implicit HRTF modeling using temporal convolutional networks. In: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Singapore: IEEE; 2021. pp. 3385-3389. DOI: 10.1109/ICASSP39728.2021.9414750
- [35] Grijalva F, Martini L, Goldenstein S, Florencio D. Anthropometric-based customization of head-related transfer functions using isomap in the horizontal plane. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). USA: IEEE; 2014. pp. 4473-4477. DOI: 10.1109/ICASSP.2014.6854448
- [36] Hu H, Zhou L, Ma H, Wu Z. HRTF personalization based on artificial neural network in individual virtual auditory space. *Applied Acoustics*. 2008;69(2): 163-172. DOI: 10.1016/j.apacoust.2007.05.007
- [37] Lee GW, Lee JH, Kim SJ, Kim HK. Directional audio rendering using a neural network based personalized HRTF. In INTERSPEECH, Brno, Czech Republic. pp. 2364–2365
- [38] Li L, Huang Q. HRTF personalization modeling based on RBF neural network. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, Canada: IEEE; 2013. pp. 3707-3710. DOI: 10.1109/ICASSP.2013.6638350
- [39] Miccini R, Spagnol S. A hybrid approach to structural modeling of individualized HRTFs. In: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). Lisbon, Portugal: IEEE; 2021. pp. 80-85. DOI: 10.1109/VRW52623.2021.00022
- [40] Shu-Nung Y, Collins T, Liang C. Head-related transfer function selection using neural networks. *Archives of Acoustics*. 2017;42(3):365-373. DOI: 10.1515/aoa-2017-0038
- [41] Zhou Y, Jiang H, Ithapu VK. On the predictability of HRTFs from ear shapes

using deep networks. In: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). London: IEEE; 2021. pp. 441-445. DOI: 10.1109/ICASSP39728.2021.9414042

[42] Bilinski P, Ahrens J, Thomas MR, Tashev IJ, Platt JC. HRTF magnitude synthesis via sparse representation of anthropometric features. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). London: IEEE; 2014. pp. 4468-4472. DOI: 10.1109/ICASSP.2014.6854447

[43] Ghorbal S, Auclair T, Soladie C, Segulier R. Pinna morphological parameters influencing HRTF sets. In: Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17). Edinburgh: University of Edinburgh; 2017. pp. 353-359

[44] Mokhtari P, Takemoto H, Nishimura R, Kato H. Vertical normal modes of human ears: Individual variation and frequency estimation from pinna anthropometry. *The Journal of the Acoustical Society of America*. 2016;**140**(2):814-831. DOI: 10.1121/1.4960481

[45] Onofrei MG, Miccini R, Unnthorsson R, Serafin S, Spagnol S. 3d ear shape as an estimator of HRTF notch frequency. In: 17th Sound and Music Computing Conference. Torino: Sound and Music Computing Network; 2020. pp. 131-137. DOI: 10.5281/zenodo.3898720

[46] Spagnol S, Geronazzo M, Avanzini F. On the relation between pinna reflection patterns and head-related transfer function features. *IEEE Transactions on Audio, Speech, and Language Processing*. 2012;**21**(3): 508-519. DOI: 10.1109/TASL.2012.2227730

[47] Pollack K, Majdak P, Furtado H. A parametric pinna model for the

calculations of head-related transfer functions. In: Proceedings of Forum Acusticum. Lyon. 2020. pp. 1357-1360. DOI: 10.48465/fa.2020.02800

[48] Stitt P, Katz BFG. Sensitivity analysis of pinna morphology on head-related transfer functions simulated via a parametric pinna model. *The Journal of the Acoustical Society of America*. 2021;**149**(4):2559-2572, ISSN: 0001-4966. DOI: 10.1121/10.0004128

[49] Katz BF, Parsehian G. Perceptually based head-related transfer function database optimization. *The Journal of the Acoustical Society of America*. 2012;**131**(2):EL99-EL105. DOI: 10.1121/1.3672641

[50] Baumgartner R, Majdak P, Laback B. Modeling sound-source localization in sagittal planes for human listeners. *The Journal of the Acoustical Society of America*. 2014;**136**(2): 791-802. DOI: 10.1121/1.4887447

[51] Xie B, Zhong X, He N. Typical data and cluster analysis on head-related transfer functions from chinese subjects. *Applied Acoustics*. 2015;**94**:1-13. DOI: 10.1016/j.apacoust.2015.01.022

[52] Toppila E, Pyykkö I, Starck J. Age and noise-induced hearing loss. *Scandinavian Audiology*. 2001;**30**(4): 236-244. DOI: 10.1080/01050390152704751

[53] Klumpp RG, Eady HR. Some measurements of interaural time difference thresholds. *The Journal of the Acoustical Society of America*. 1956;**28**: 859-860. DOI: 10.1121/1.1908493

[54] Blauert J. Spatial hearing. In: *The Psychophysics of Human Sound Localization*. Cambridge, MA: The MIT Press; 1997

[55] Raykar VC, Duraiswami R, Yegnanarayana B. Extracting the frequencies of the pinna spectral

notches in measured head related impulse responses. *The Journal of the Acoustical Society of America*. 2005; **118**(1):364-374. DOI: 10.1121/1.1923368

[56] Takemoto H, Mokhtari P, Kato H, Nishimura R, Iida K. Mechanism for generating peaks and notches of head-related transfer functions in the median plane. *The Journal of the Acoustical Society of America*. 2012;**132**(6): 3832-3841. DOI: 10.1121/1.4765083

[57] Algazi VR, Duda RO, Duraiswami R, Gumerov NA, Tang Z. Approximating the head-related transfer function using simple geometric models of the head and torso. *The Journal of the Acoustical Society of America*. 2002;**112**(5): 2053-2064. DOI: 10.1121/1.1508780

[58] Macpherson EA, Middlebrooks JC. Vertical-plane sound localization probed with ripple-spectrum noise. *The Journal of the Acoustical Society of America*. 2003;**114**(1):430-445. DOI: 10.1121/1.1582174

[59] Goupell MJ, Majdak P, Laback B. Median-plane sound localization as a function of the number of spectral channels using a channel vocoder. *The Journal of the Acoustical Society of America*. 2010;**127**(2):990-1001. DOI: 10.1121/1.3283014

[60] Kulkarni A, Colburn HS. Role of spectral detail in sound-source localization. *Nature*. 1998;**396**(6713): 747-749. DOI: 10.1038/25526

[61] Senova MA, McAnally KI, Martin RL. Localization of virtual sound as a function of head-related impulse response duration. *Journal of the Audio Engineering Society*. 2002;**50**(1/2):57-66

[62] Thavam S, Dietz M. Smallest perceivable interaural time differences. *The Journal of the Acoustical Society of America*. 2019;**145**(1):458-468. DOI: 10.1121/1.5087566

[63] Andreopoulou A, Katz BF. Identification of perceptually relevant methods of inter-aural time difference estimation. *The Journal of the Acoustical Society of America*. 2017; **142**(2):588-598. DOI: 10.1121/1.4996457

[64] Katz BF, Noisternig M. A comparative study of interaural time delay estimation methods. *The Journal of the Acoustical Society of America*. 2014;**135**(6):3530-3540. DOI: 10.1121/1.4875714

[65] Algazi R, Avendano C, Duda RO. Estimation of a spherical-head model from anthropometry. *Journal of the Audio Engineering Society*. 2001;**49**: 472-479

[66] Zhang W, Abhayapala TD, Kennedy RA, Duraiswami R. Insights into head-related transfer function: Spatial dimensionality and continuous representation. *The Journal of the Acoustical Society of America*. 2010; **127**(4):2347-2357. DOI: 10.1121/1.3336399

[67] Bomhardt R, de la Fuente Klein M, Fels J. A high-resolution head-related transfer function and three-dimensional ear model database. In: *Proceedings of Meetings on Acoustics 172ASA*. Vol. 29. Illinois, United States: ASA; 2016. p. 050002. DOI: 10.1121/2.0000467

[68] Carpentier T, Bahu H, Noisternig M, Warusfel O. Measurement of a head-related transfer function database with high spatial resolution. In: *7th Forum Acusticum (EAA)*. Ukraine: EAA; 2014

[69] Jin CT, Guillon P, Epain N, Zolfaghari R, Van Schaik A, Tew AI, et al. Creating the Sydney York morphological and acoustic recordings of ears database. *IEEE Transactions on Multimedia*. 2013;**16**(1):37-46. DOI: 10.1109/TMM.2013.2282134

[70] Mills AW. On the minimum audible angle. *The Journal of the Acoustical*



- Society of America. 1958;**30**(4):237-246. DOI: 10.1121/1.1909553
- [71] Wersényi G. HRTFs in human localization: Measurement, spectral evaluation and practical use in virtual audio environment. Dissertation. Cottbus, Germany: Brandenburg University of Technology; 2002
- [72] Zhong X, Xie B, et al. Head-related transfer functions and virtual auditory display. In: *Soundscape Semiotics-Localization and Categorization*. Plantation, FL, United States: J. Ross Publishing; 2014. p. 1. DOI: 10.5772/56907
- [73] Makous JC, Middlebrooks JC. Two-dimensional sound localization by human listeners. *The Journal of the Acoustical Society of America*. 1990; **87**(5):2188-2200. DOI: 10.1121/1.399186
- [74] Middlebrooks JC. Spectral shape cues for sound localization. In: *Binaural and Spatial Hearing in Real and Virtual Environments*. New York: Psychology Press; 1997. pp. 77-97
- [75] Middlebrooks JC. Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. *The Journal of the Acoustical Society of America*. 1999; **106**(3):1493-1510. DOI: 10.1121/1.427147
- [76] Perrott DR, Saberi K. Minimum audible angle thresholds for sources varying in both elevation and azimuth. *The Journal of the Acoustical Society of America*. 1990;**87**(4):1728-1731, ISSN: 0001-4966. DOI: 10.1121/1.399421
- [77] Middlebrooks JC, Green DM. Sound localization by human listeners. *Annual Review of Psychology*. 1991;**42**(1): 135-159. DOI: 10.1146/annurev.ps.42.020191.001031
- [78] Poirier P, Miljours S, Lassonde M, Lepore F. Sound localization in acallosal human listeners. *Brain*. 1993;**116**(1): 53-69. DOI: 10.1093/brain/116.1.53
- [79] Voss P, Lassonde M, Gougoux F, Fortin M, Guillemot J-P, Lepore F. Early- and late-onset blind individuals show supra-normal auditory abilities in far-space. *Current Biology*. 2004; **14**(19):1734-1738. DOI: 10.1016/j.cub.2004.09.051
- [80] Senn P, Kompis M, Vischer M, Haeusler R. Minimum audible angle, just noticeable interaural differences and speech intelligibility with bilateral cochlear implants using clinical speech processors. *Audiology and Neurotology*. 2005;**10**(6):342-352. DOI: 10.1159/000087351
- [81] Pulkki V. Localization of amplitude-panned virtual sources II: Two- and three-dimensional panning. *Journal of the Audio Engineering Society*. 2001; **49**(4):753-767
- [82] Bremen P, van Wanrooij MM, van Opstal AJ. Pinna cues determine orienting response modes to synchronous sounds in elevation. *Journal of Neuroscience*. 2010;**30**(1): 194-204. DOI: 10.1523/JNEUROSCI.2982-09.2010
- [83] Brimijoin WO, Akeroyd MA. The moving minimum audible angle is smaller during self motion than during source motion. *Frontiers in Neuroscience*. 2014;**8**:273. DOI: 10.3389/fnins.2014.00273
- [84] Begault DR, Wenzel EM, Anderson MR. Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *Journal of the Audio Engineering Society*. 2001; **49**(10):904-916
- [85] Stitt P, Hendrickx E, Messonnier J, Katz B. The role of head tracking in binaural rendering. In: 29th

- Tonmeistertagung, International VDT Convention. Germany: CCN Cologne; 2016
- [86] Urbanietz C, Enzner G. Binaural rendering of dynamic head and sound source orientation using high-resolution HRTF and retarded time. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary, AB, Canada: IEEE; 2018. pp. 566-570. DOI: 10.1109/ICASSP.2018.8461343
- [87] Pörschmann C, Arend JM. Obtaining dense HRTF sets from sparse measurements in reverberant environments. In: Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio. New York, New York, United States: Audio Engineering Society; 2019
- [88] Algazi VR, Duda RO, Thompson DM, Avendano C. The CIPIC HRTF database. In: Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575). New York: IEEE; 2001. pp. 99-102. DOI: 10.1109/ASPAA.2001.9695520
- [89] Pelzer R, Dinakaran M, Brinkmann F, Lepa S, Grosche P, Weinzierl S. Head-related transfer function recommendation based on perceptual similarities and anthropometric features. *The Journal of the Acoustical Society of America*. 2020; **148**(6):3809-3817. DOI: 10.1121/10.0002884
- [90] Ziegelwanger H, Reichinger A, Majdak P. Calculation of listener-specific head-related transfer functions: Effect of mesh quality. In: Proceedings of Meetings on Acoustics. Vol. 19. Montreal, Canada. 2013. p. 050017. DOI: 10.1121/1.4799868
- [91] Gardner MB, Gardner RS. Problem of localization in the median plane: Effect of pinnae cavity occlusion. *The Journal of the Acoustical Society of America*. 1973;**53**(2):400-408. DOI: 10.1121/1.1913336
- [92] Nelson PA, Kahana Y. Spherical harmonics, singular-value decomposition and head-related transfer function. *Journal of Sound and Vibration*. 2001;**239**:607-637. DOI: 10.1006/jsvi.2000.3227
- [93] Shaw EAG. The external ear. In: Keidel WD, Neff WD, editors. *Auditory System*. Vol. 5/1. Berlin Heidelberg, ISBN: 978-3-642-65831-0 978-3-642-65829-7: Springer; 1974. pp. 455-490. DOI: 10.1007/978-3-642-65829-7\_14
- [94] Brinkmann F. The FABIAN head-related transfer function data base. Berlin: Technische Universität Berlin; 2017. DOI: 10.14279/depositonce-5718
- [95] Brinkmann F, Dinakaran M, Pelzer R, Grosche P, Voss D, Weinzierl S. A cross-evaluated database of measured and simulated HRTFs including 3D head meshes, anthropometric features, and headphone impulse responses. *Journal of the Audio Engineering Society*. 2019; **67**(9):705-718. DOI: 10.17743/jaes.2019.0024
- [96] Ghorbal S, Bonjour X, Séguier R. Computed HRIRs and ears database for acoustic research. In: Audio Engineering Society Convention 148. New York, New York, United States: Audio Engineering Society; 2020
- [97] Katz BF. Acoustic absorption measurement of human hair and skin within the audible frequency range. *The Journal of the Acoustical Society of America*. 2000;**108**(5 Pt 1):2238-2242. DOI: 10.1121/1.1314319
- [98] Treeby BE, Pan J, Paurobally RM. An experimental study of the acoustic impedance characteristics of human

hair. *The Journal of the Acoustical Society of America*. 2007;**122**(4): 2107-2117. DOI: 10.1121/1.2773946

[99] Brinkmann F, Lindau A, Weinzierl S. On the authenticity of individual dynamic binaural synthesis. *The Journal of the Acoustical Society of America*. 2017;**142**(4):1784-1795, ISSN: 0001-4966. DOI: 10.1121/1.5005606

[100] Brinkmann F, Lindau A, Weinzierl S, Müller-Trapet M, Opdam R, Vorländer M, et al. A high resolution and full-spherical head-related transfer function database for different head-above-torso orientations. *Journal of the Audio Engineering Society*. 2017;**65**(10):841-848. DOI: 10.17743/jaes.2017.0033

[101] Ziegelwanger H, Majdak P, Kreuzer W. Numerical calculation of listener-specific head-related transfer functions and sound localization: Microphone model and mesh discretization. *The Journal of the Acoustical Society of America*. 2015; **138**(1):208-222, ISSN: 0001-4966. DOI: 10.1121/1.4922518

[102] Zotkin DN, Duraiswami R, Grassi E, Gumerov NA. Fast head-related transfer function measurement via reciprocity. *The Journal of the Acoustical Society of America*. 2006;**120**(4):2202-2215. DOI: 10.1121/1.2207578

[103] Carlile S, Leong P, Hyams S. The nature and distribution of errors in sound localization by human listeners. *Hearing Research*. 1997;**114**(1-2): 179-196. DOI: 10.1016/S0378-5955(97)00161-5

[104] Masiero B, Pollow M, Fels J. Design of a fast broadband individual head-related transfer function measurement system. Vol. 97. Hirzel: Acustica; 2011. pp. 136-136

[105] Bau D, Lübeck T, Arend JM, Dziwis D, Pörschmann C. Simplifying

head-related transfer function measurements: A system for use in regular rooms based on free head movements. In: 8th International Conference of Immersive and 3D Audio. Bologna, Italy: I3DA; 2021

[106] Reijniers J, Partoens B, Steckel J, Peremans H. HRTF measurement by means of unsupervised head movements with respect to a single fixed speaker. Vol. 8. London: IEEE Access; 2020. pp. 92287-92300, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.2994932

[107] Fukudome K, Suetsugu T, Ueshin T, Idegami R, Takeya K. The fast measurement of head related impulse responses for all azimuthal directions using the continuous measurement method with a servo-swiveled chair. *Applied Acoustics*. 2007;**68**(8):864-884. DOI: 10.1016/j.apacoust.2006.09.009

[108] He J, Ranjan R, Gan W-S, Chaudhary NK, Hai ND, Gupta R. Fast continuous measurement of HRTFs with unconstrained head movements for 3d audio. *Journal of the Audio Engineering Society*. 2018;**66**(11): 884-900. DOI: 10.17743/jaes.2018.0050

[109] Richter J-G, Fels J. On the influence of continuous subject rotation during high-resolution head-related transfer function measurements. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 2019;**27**(4): 730-741. DOI: 10.1109/TASLP.2019.2894329

[110] Pulkki V, Laitinen M-V, Sivonen V. HRTF measurements with a continuously moving loudspeaker and swept sines. In: *Audio Engineering Society Convention 128*. New York, New York, United States: Audio Engineering Society; 2010

[111] Kabzinski T, Jax P. Towards faster continuous multi-channel HRTF measurements based on learning system



- models. In: 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Singapore: IEEE; 2021 arXiv preprint arXiv: 2110.03630
- [112] Majdak P, Balazs P, Laback B. Multiple exponential sweep method for fast measurement of head-related transfer functions. *Journal of the Audio Engineering Society*. 2007;**55**:623-637
- [113] Middlebrooks JC, Makous JC, Green DM. Directional sensitivity of sound-pressure levels in the human ear canal. *The Journal of the Acoustical Society of America*. 1989;**86**(1):89-108. DOI: 10.1121/1.398224
- [114] Wightman F, Kistler D, Foster S, Abel J. A comparison of head-related transfer functions measured deep in the ear canal and at the ear canal entrance. In: 17th Midwinter Meeting of the Association for Research in Otolaryngology. Vol. 71. Montreal: ARO; 1995
- [115] Zahorik P. Limitations in using golog codes for head-related transfer function measurement. *The Journal of the Acoustical Society of America*. 2000; **107**(3):1793-1796. DOI: 10.1121/1.428579
- [116] Dietrich P, Masiero B, Vorländer M. On the optimization of the multiple exponential sweep method. *Journal of the Audio Engineering Society*. 2013;**61**(3):113-124
- [117] Armstrong C, Thresh L, Murphy D, Kearney G. A perceptual evaluation of individual and non-individual HRTFs: A case study of the SADIE II database. *Applied Sciences*. 2018;**8**(11):2029. DOI: 10.3390/app8112029
- [118] Denk F, Kollmeier B, Ewert SD. Removing reflections in semianechoic impulse responses by frequency-dependent truncation. *Journal of the Audio Engineering Society*. 2018;**66**(3): 146-153. DOI: 10.17743/jaes.2018.0002
- [119] Kistler DJ, Wightman FL. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *The Journal of the Acoustical Society of America*. 1992;**91**(3):1637-1647. DOI: 10.1121/1.402444
- [120] Kohlrausch A, Breebaart J. Perceptual (ir) relevance of HRTF magnitude and phase spectra. In: *Audio Engineering Society Convention 110*. New York, New York, United States: Audio Engineering Society; 2001
- [121] Bergman DR. *Computational Acoustics: Theory and Implementation*. Hoboken, New Jersey, United States: John Wiley & Sons; 2018
- [122] Marburg S. Six boundary elements per wavelength. Is that enough? *Journal of Computational Acoustics*. 2002;**10**: 25-51. DOI: 10.1142/S0218396X02001401
- [123] Botsch M, Kobbelt L. A remeshing approach to multiresolution modeling. In: *Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*. New York, NY, United States: Association for Computing Machinery; 2004. pp. 185-192. DOI: 10.1145/1057432.1057457
- [124] Reichinger A, Majdak P, Sablatnig R, Maierhofer S. Evaluation of methods for optical 3-D scanning of human pinnae. In: *Proceedings of the 3D Vision Conference*. Seattle, WA: IEEE; 2013. pp. 390-397. DOI: 10.1109/3DV.2013.58
- [125] Dinakaran M, Brinkmann F, Harder S, Pelzer R, Grosche P, Paulsen RR, et al. Perceptually motivated analysis of numerically simulated head-related transfer functions generated by various 3d surface scanning systems. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

Calgary, Alberta, Canada: IEEE; 2018.  
pp. 551-555. DOI: 10.1109/  
ICASSP.2018.8461789

[126] Greff R, Katz BF. Round robin comparison of HRTF simulation systems: Preliminary results. In: Audio Engineering Society Convention 123, Page Convention Paper 7188. New York, New York, United States: Audio Engineering Society; 2007

[127] Dellepiane M, Pietroni N, Tsingos N, Asselot M, Scopigno R. Reconstructing head models from photographs for individualized 3d-audio processing. In: Computer Graphics Forum. Vol. 27. Hoboken, New Jersey, United States: Wiley Online Library; 2008. pp. 1719-1727. DOI: 10.1111/j.1467-8659.2008.01316.x

[128] Iida K, Nishiyama O, Aizaki T. Estimation of the category of notch frequency bins of the individual head-related transfer functions using the anthropometry of the listener's pinnae. *Applied Acoustics*. 2021;177:107929. DOI: 10.1016/j.apacoust.2021.107929

[129] Pollack K, Brinkmann F, Majdak P, Kreuzer W. Von Fotos zu personalisierter räumlicher Audiowiedergabe [from photos to personalised spatial audio playback]. *e & i Elektrotechnik und Informationstechnik*. 2021;138(3):1-6. DOI: 10.1007/s00502-021-00891-4

[130] Ullman S, Brenner S. The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B. Biological Sciences*. 1979;203(1153):405-426. DOI: 10.1098/rspb.1979.0006 Publisher: Royal Society

[131] Sommerfeld A. *Partial Differential Equations in Physics*. Cambridge, Massachusetts, United States: Academic Press; 1949

[132] Turner MJ, Clough RW, Martin HC, Topp L. Stiffness and deflection analysis of complex

structures. *Journal of the Aeronautical Sciences*. 1956;23(9):805-823. DOI: 10.2514/8.3664

[133] Bériot H, Prinn A, Gabard G. Efficient implementation of high-order finite elements for Helmholtz problems. *International Journal for Numerical Methods in Engineering*. 2016;106(3): 213-240. DOI: 10.1002/nme.5172

[134] Gabard G, Bériot H, Prinn A, Kucukcoskun K. Adaptive, high-order finite-element method for convected acoustics. *AIAA Journal*. 2018;56(8): 3179-3191. DOI: 10.2514/1.J057054

[135] Ueberhuber CW. *Numerical Computation 1: Methods, Software, and Analysis*. Vol. 16. Berlin, Germany: Springer Science & Business Media; 1997

[136] Bériot H, Modave A. An automatic perfectly matched layer for acoustic finite element simulations in convex domains of general shape. *International Journal for Numerical Methods in Engineering*. 2021;122(5):1239-1261. DOI: 10.1002/nme.6560

[137] Farahikia M, Su QT. Optimized finite element method for acoustic scattering analysis with application to head-related transfer function estimation. *Journal of Vibration and Acoustics*. 2017;139(3):034501. DOI: 10.1115/1.4035813

[138] Harder S, Paulsen RR, Larsen M, Laugesen S, Mihocic M, Majdak P. A framework for geometry acquisition, 3-D printing, simulation, and measurement of head-related transfer functions with a focus on hearing-assistive devices. *Computer Aided Design*. 2016;75-76:39-46, ISSN: 0010-4485. DOI: 10.1016/j.cad.2016.02.006

[139] Huttunen T, Seppälä ET, Kirkeby O, Kärkkäinen A, Kärkkäinen L. Simulation of the transfer function for a head-and-torso model

- over the entire audible frequency range. *Journal of Computational Acoustics*. 2007;**15**(04):429-448. DOI: 10.1142/S0218396X07003469
- [140] Kahana Y. *Numerical Modelling of the Head-Related Transfer Function*. Southampton, UK: University of Southampton; 2000
- [141] Ma F, Wu JH, Huang M, Zhang W, Hou W, Bai C. Finite element determination of the head-related transfer function. *Journal of Mechanics in Medicine and Biology*. 2015;**15**(05): 1550066. DOI: 10.1142/S0219519415500669
- [142] Yee K. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*. 1966;**14**(3):302-307. DOI: 10.1109/TAP.1966.1138693
- [143] Botts J, Savioja L. Spectral and pseudospectral properties of finite difference models used in audio and room acoustics. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 2014;**22**(9): 1403-1412. DOI: 10.1109/TASLP.2014.2332045
- [144] Häggblad J, Runborg O. Accuracy of staircase approximations in finite-difference methods for wave propagation. *Numerische Mathematik*. 2014;**128**(4):741-771. DOI: 10.1007/s00211-014-0625-1
- [145] Prepelitã ST, Geronazzo M, Avanzini F, Savioja L. Influence of voxelization on finite difference time domain simulations of head-related transfer functions. *The Journal of the Acoustical Society of America*. 2016; **139**(5):2489-2504. DOI: 10.1121/1.4947546
- [146] Prepelitã ST, Gómez Bolaños J, Geronazzo M, Mehra R, Savioja L. Pinna-related transfer functions and lossless wave equation using finite-difference methods: Verification and asymptotic solution. *The Journal of the Acoustical Society of America*. 2019; **146**(5):3629-3645. DOI: 10.1121/1.5131245
- [147] Prepelitã ST, Gómez Bolaños J, Geronazzo M, Mehra R, Savioja L. Pinna-related transfer functions and lossless wave equation using finite-difference methods: Validation with measurements. *The Journal of the Acoustical Society of America*. 2020; **147**(5):3631-3645. DOI: 10.1121/10.0001230
- [148] Botteldooren D. Acoustical finite-difference time-domain simulation in a quasi-cartesian grid. *The Journal of the Acoustical Society of America*. 1994; **95**(5):2313-2319. DOI: 10.1121/1.409866
- [149] Willemsen S, Bilbao S, Ducceschi M, Serafin S. Dynamic grids for finite-difference schemes in musical instrument simulations. In: *24th International Conference on Digital Audio Effects*. Vienna, Austria: DAFX; 2021. pp. 144-151
- [150] Bilbao S. Modeling of complex geometries and boundary conditions in finite difference/finite volume time domain room acoustics simulation. *IEEE Transactions on Audio, Speech, and Language Processing*. 2013;**21**(7): 1524-1533. DOI: 10.1109/TASL.2013.2256897
- [151] Bilbao S, Hamilton B. Passive volumetric time domain simulation for room acoustics applications. *The Journal of the Acoustical Society of America*. 2019;**145**(4):2613-2624. DOI: 10.1121/1.5095876
- [152] Bilbao S, Hamilton B, Botts J, Savioja L. Finite volume time domain room acoustics simulation under general impedance boundary conditions. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 2015;**24**(1):



161-173. DOI: 10.1109/  
TASLP.2015.25000180

[153] Peiró, J. Sherwin S. Finite difference, finite element and finite volume methods for partial differential equations. In *Handbook of Materials Modeling*. Berlin, Germany: Springer; 2005. pp. 2415–2446. DOI: 10.1007/978-1-4020-3286-8\_127

[154] Mokhtari P, Takemoto H, Nishimura R, Kato H. Frequency and amplitude estimation of the first peak of head-related transfer functions from individual pinna anthropometry. *The Journal of the Acoustical Society of America*. 2015;**137**(2):690-701. DOI: 10.1121/1.4906160

[155] Xiao T, Huo Liu Q. Finite difference computation of head-related transfer function for human hearing. *The Journal of the Acoustical Society of America*. 2003;**113**(5):2434-2441, ISSN: 0001-4966. DOI: 10.1121/1.1561495

[156] Gumerov NA, O'Donovan AE, Duraiswami R, Zotkin DN. Computation of the head-related transfer function via the fast multipole accelerated boundary element method and its spherical harmonic representation. *The Journal of the Acoustical Society of America*. 2010; **127**(1):370-386. DOI: 10.1121/1.3257598

[157] Galerkin BG. Rods and plates. Series occurring in various questions concerning the elastic equilibrium of rods and plates. *Engineers Bulletin (Vestnik Inzhenerov)*. 1915;**19**:897-908

[158] Nyström EJ. Über die praktische Auflösung von Integralgleichungen mit Anwendungen auf Randwertaufgaben [about the practical solution of integral equations with applications to boundary value problems]. *Acta Mathematica*. 1930; **54**:185-204. DOI: 10.1007/BF02547521

[159] Sauter S, Schwab S. *Boundary Element Methods*. Berlin, Germany: Springer; 2011

[160] Arnold DN, Wendland WL. Collocation versus Galerkin procedures for boundary integral methods. In: Brebbia CA, editor. *Boundary Element Methods in Engineering*. Berlin, Germany ISBN: 978-3-662-11275-5: Springer International Publishing; 1982. DOI: 10.1007/978-3-662-11273-1\_2

[161] Duffy MG. Quadrature over a pyramid or cube of integrands with a singularity at a vertex. *SIAM Journal on Numerical Analysis*. 1982;**19**(6): 1260-1262. DOI: 10.1137/0719090

[162] Krishnasamy G, Schmerr L, Rudolph T, Rizzo F. Hypersingular boundary integral equations: Some applications in acoustic and elastic wave scattering. *Transactions of the ASME*. 1990;**57**:404-414. DOI: 10.1115/1.2892004

[163] Coifman R, Rokhlin V, Wandzura S. The fast multipole method for the wave equations: A pedestrian prescription. *IEEE Antennas and Propagation Magazine*. 1993;**35**(3):7-12, ISSN: 1045-9243. DOI: 10.1109/74.250128

[164] Hackbusch W. *Hierarchical Matrices: Algorithms and Analysis*. Berlin, Heidelberg: Springer; 2015. DOI: 10.1007/978-3-662-47324-5

[165] Kreuzer W, Majdak P, Chen Z. Fast multipole boundary element method to calculate head-related transfer functions for a wide frequency range. *The Journal of the Acoustical Society of America*. 2009;**126**(3):1280-1290. DOI: 10.1121/1.3177264

[166] Saad Y. *Iterative Methods for Sparse Linear Systems*. New Delhi, India: SIAM; 2003

[167] Burton AJ, Miller GF. The application of integral equation methods to the numerical solution of some exterior boundary-value problems. *Proceedings of the Royal Society of*

London A. *Mathematical and Physical Sciences*. 1971;**323**(1553):201-210, ISSN: 0080-4630. DOI: 10.1098/rspa.1971.0097

[168] Katz BF. Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation. *The Journal of the Acoustical Society of America*. 2001;**110** (5 Pt 1):2440-2448. DOI: 10.1121/1.1412440

[169] Katz BF. Boundary element method calculation of individual head-related transfer function. II. Impedance effects and comparisons to real measurements. *The Journal of the Acoustical Society of America*. 2001;**110** (5 Pt 1):2449-2455. DOI: 10.1121/1.1412441

[170] Otani M, Ise S. A fast calculation method of the head-related transfer functions for multiple source points based on the boundary element method. *Acoustical Science and Technology*. 2003;**24**(5):259-266. DOI: 10.1250/ast.24.259

[171] Otani M, Ise S. Fast calculation system specialized for head-related transfer function based on boundary element method. *The Journal of the Acoustical Society of America*. 2006;**119** (5 Pt 1):2589-2598, ISSN: 0001-4966. DOI: 10.1121/1.2191608

[172] Ziegelwanger H, Kreuzer W, Majdak P. Mesh2HRTF: Open-source software package for the numerical calculation of head-related transfer functions. In *Proceedings of the 22nd International Congress on Sound and Vibration*, 1–8, IEEE Florence, IT. 2015. DOI: 10.13140/RG.2.1.1707.1128

[173] Fink KJ, Ray L. Individualization of head related transfer functions using principal component analysis. *Applied Acoustics*. 2015;**87**: 162-173. DOI: 10.1016/j.apacoust.2014.07.005

[174] Xie B, Zhong X, Rao D, Liang Z. Head-related transfer function database and its analyses. *Science in China Series G: Physics, Mechanics and Astronomy*. 2007;**50**(3):267-280, ISSN: 1672-1799, 1862-2844. DOI: 10.1007/s11433-007-0018-x

[175] Nishino T, Inoue N, Takeda K, Itakura F. Estimation of HRTFs on the horizontal plane using physical features. *Applied Acoustics*. 2007;**68**(8):897-908, ISSN: 0003-682X. DOI: 10/dr4tg3

[176] Xie B. *Head-Related Transfer Function and Virtual Auditory Display*. Plantation, FL, United States: J. Ross Publishing; 2013

[177] Gromov M. Metric structures for Riemannian and non-Riemannian spaces. *Bulletin of the American Mathematical Society*. 2001;**38**:353-363

[178] Hebrank J, Wright D. Are two ears necessary for localization of sound sources on the median plane? *The Journal of the Acoustical Society of America*. 1974;**56**(3):935-938. DOI: 10.1121/1.1903351