Chapter 1
Introduction

Abstract This introductory chapter describes the motivation for writing this book. Then, head-related transfer functions (HRTFs) are defined, and the coordinate system used in this book is described. Finally, recent advancements in HRTF research are briefly summarized.

1.1 Why Research HRTFs

Humans perceive the direction and distance of a sound source in three-dimensional space with only two ears. Geometrically, n + 1 receiving points are required to identify the position of an object in n-dimensional space. Limiting the discussion to the direction of sound, it is thus impossible to identify the direction geodetically with two receiving points (two ears). So, how do we perceive the direction of sound? This is the starting point of HRTF research.

An easy way to identify the direction of sound is to rotate your head. If you rotate your head so that you perceive a sound as being in front of you, you are now looking in the direction of the sound source. However, humans can perceive the direction of sound without rotating their heads, and in fact rarely rotate their heads spontaneously to identify the direction of sound (Nojima et al. 2013). Even barn owls, which catch prey by identifying the direction of the sound, do not use head movement in performing this task (Knudsen and Konishi 1979).

HRTFs play a key role in the perception of sound direction. With few exceptions (e.g., listening to music through headphones and telephone conversations), most of the sounds we hear in everyday life are affected by the HRTF. Therefore, research on HRTFs is absolutely essential to understanding the mechanism behind the perception of sound direction.

On the other hand, the reproduction of an existing sound field and/or the creation of a virtual sound field beyond time and space, i.e., acoustic virtual reality (VR) and augmented reality (AR), can be achieved through the use of HRTFs.

The original version of this chapter was revised. The correction to this chapter is available at https://doi.org/10.1007/978-981-13-9745-5_14

K. Iida, *Head-Related Transfer Function and Acoustic Virtual Reality*,
https://doi.org/10.1007/978-981-13-9745-5_1

As such, HRTFs play an important role in both research on spatial hearing and development of acoustical VR systems. So, what exactly is an HRTF?

1.2 What Is an HRTF?

Sound waves are affected by the head, pinnae, and torso of the listener before reaching the listener's ear drums. The HRTF is the physical change in the sound wave expressed in the frequency domain due to these effects. Figure 1.1 shows an amplitude spectrum of an HRTF for a sound source located in front of a subject. Here, 0 dB indicates the amplitude of the sound pressure measured without the head, pinnae, and torso. There exist several distinguishing peaks and notches that exceed ±10 dB. A positive value of the amplitude of the HRTF indicates that the sound pressure increases as a result of the head, pinnae, and torso of the subject, whereas a negative value indicates that the sound pressure decreases as a result of these components. Humans hear such a sound with spectral peaks and notches in everyday life.

The HRTF is defined as follows:

$$H_{1,\mathrm{r}}(s,\alpha,\beta,r,\omega)=\frac{G_{1,\mathrm{r}}(s,\alpha,\beta,r,\omega)}{F(\alpha,\beta,r,\omega)} \tag{1.1}$$

where $G_{l,r}$ is the transfer function between a sound source and the entrance of the ear canal[1] of the subject in the free field, F is the transfer function between a sound

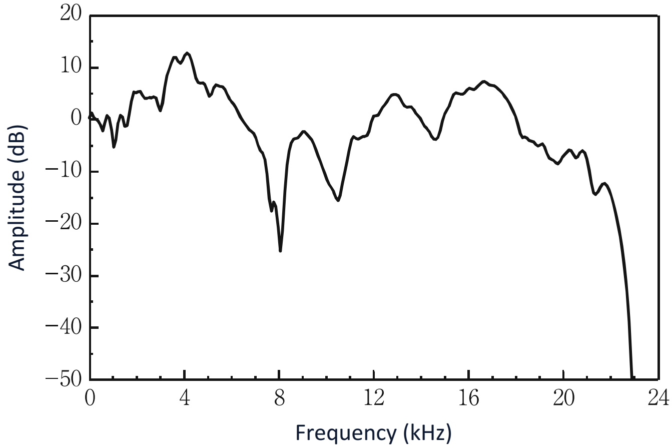

Fig. 1.1 Example of the amplitude of an HRTF for the forward direction

[1]There is another way to define the HRTF, in which the receiving point is the ear drum rather than the entrance of the ear canal (Wightman and Kistler 1989a, b). However, it is not easy to place a miniature microphone just before the ear drum to measure the HRTF. Moreover, the transfer function of the ear canal does not depend on the direction of the sound source (Møller 1992). As such, most of studies measure HRTFs using a miniature microphone located at the entrance of the ear canal.

source and the point corresponding to the center of the subject's head in the free field without the subject. The subscripts l and r denote the left and right ears, and s denotes the subject. Moreover, α, β, and γ denote the lateral angle, the vertical (rising) angle, and the distance, respectively. However, the distance of a sound source does not affect the HRTF when the distance exceeds 1 m (Morimoto et al. 1976).

Note that the phase spectrum of an HRTF is not described in this book because the absolute phase of the HRTF does not play an important role in the perception of sound direction (Kulkarni et al. 1999) if only the relative phase difference between the left and right ears is preserved.

The HRTF varies with the direction of a sound source. This is due to the asymmetrical shape of the head and pinnae. Humans perceive the direction of sound using the directional dependence of HRTF. The directional dependence of HRTF is described in detail in Chaps. 2 and 3.

The HRTF also varies with the listener, even if the direction of a sound source is constant, due to individual differences in shape of the head and pinna. The individualized nature of the HRTF is a serious problem in the design of practical acoustic VR systems. The details of the problem and some approaches to solving this problem are presented in Chap. 4.

1.3 HRTF and HRIR

As described above, the HRTF expresses the change in sound pressure due to the head, pinna, and torso in the frequency domain. However, the expression in the time domain may sometime be better for clarifying the characteristics of this change. The expression in the time domain is called the head-related impulse response (HRIR).

In most situations, the HRTF is obtained by Fourier transform of the measured HRIR, as follows:

$$H_{\mathrm{l,r}}(s,\alpha,\beta,r,\omega) = \frac{\mathcal{F}\left[g_{\mathrm{l,r}}(s,\alpha,\beta,r,t)\right]}{\mathcal{F}[f(\alpha,\beta,r,t)]} \tag{1.2}$$

where $g_{l,r}$ is the impulse response between a sound source and the entrance of the ear canal of the subject in the free field, f is the impulse response between a sound source and the point corresponding to the center of the subject's head in the free field without a subject. Moreover, $\mathcal{F}$ denotes the Fourier transform.

Figure 1.2 shows an HRIR for a sound source located in front of a subject. The response converges within approximately 2–3 ms.

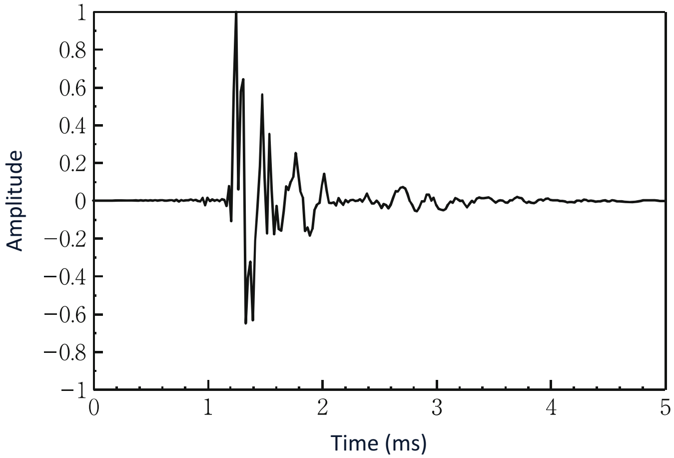

Fig. 1.2 Example of an HRIR for the forward direction

1.4 Sound Source and Sound Image

A sound wave causes many kinds of sensations when it arrives at the ear drum of a listener. What is perceived auditorily is called a sound image, or an auditory event. Whereas a sound source is a physical presence, a sound image is a psychological presence (perceptual phenomenon). A sound image includes temporal characteristics (sense of reverberation, rhythm, duration, and so on), spatial characteristics (sense of direction, distance, broadening, and so on), and qualitative characteristics (loudness, pitch, timbre, and so on). Perception of the direction and distance of a sound image is called sound image localization.

In most situations in our everyday life, a sound image is localized at the position of a sound source. However, this is not always the case. For a narrow-band signal, a sound image is sometimes localized to the front for a sound source that is actually located to the rear, and vice versa. This phenomenon is referred to as front-back confusion. Moreover, for certain narrow-band frequency ranges, a sound image is often localized in a particular direction, regardless of the direction of the actual sound source (Blauert 1969/70). These phenomena are discussed in detail in Chaps. 3, 6, and 7.

On the other hand, acoustic VR and AR systems intentionally make the listener perceive a sound image at an arbitrary position in three-dimensional space, while the sound sources (headphones) actually exist near the listener's ears.

1.5 Coordinate System

In this section, I explain the coordinate system used in this book. Figure 1.3 shows an illustration of a spherical coordinate system, which uses azimuth angle ϕ and elevation angle θ. This coordinate system is likely familiar to most readers. The azimuth angle and the elevation angle, however, do not correspond to the human mechanisms of perception of sound direction.

Thus, the interaural-polar-axis coordinate system is used in this book (Fig. 1.4). In this coordinate system, the direction of a sound source is expressed in terms of lateral angle α and vertical (rising) angle β. The lateral angle α corresponds to the mechanism for left-right perception (lateral direction) of a sound image, and the vertical angle β corresponds to that for front-back and up-down perception (vertical direction) of a sound image.

The lateral angle α is the complement of the angle between the aural axis and the line segment connecting the sound source and the origin. The origin is the center of the subject's head, and the aural axis is the straight line connecting the entrances of the left and right ear canals and the origin. The vertical angle β is the elevation within the sagittal plane, which passes through a sound source.

The horizontal plane is the plane which passes the right eye orbit and the left and right tragi (Frankfurt horizontal plane). The transverse plane is the plane perpendicular to the horizontal plane that passes through the entrances of the left and right ear canals. The median plane is the plane perpendicular to both the horizontal plane and

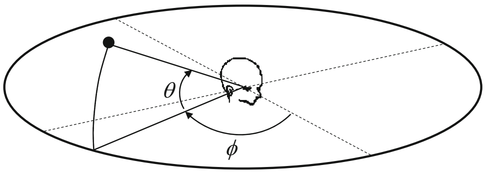

Fig. 1.3 Spherical coordinate system

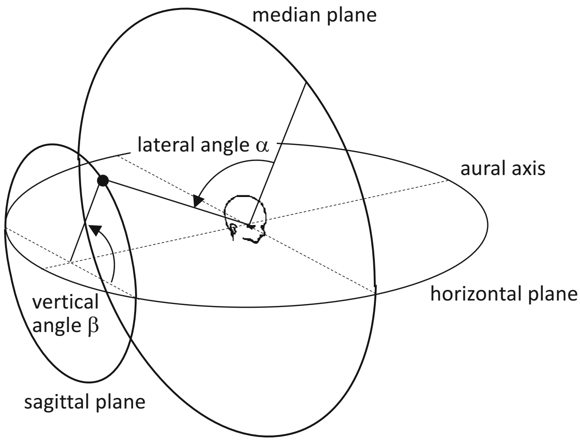

Fig. 1.4 Interaural-polar-axis coordinate system

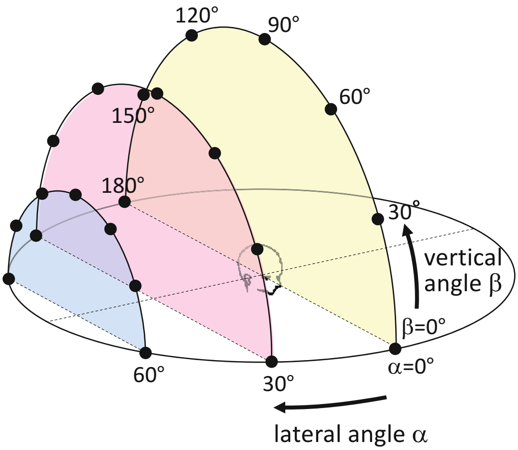

Fig. 1.5 Examples of points in sagittal planes

the transverse plane that divides the body exactly into left and right sides. The sagittal planes are the planes parallel to the median plane.

Figure 1.5 shows examples of the points, the lateral angles of which are 0°, 30°, and 60° and the vertical angles of which are 0° to 180°, in 30° steps. The interaural-polar-axis coordinate system defines the direction of a sound source using two pieces of information: the lateral sagittal plane on which the sound source is located and the elevation angle in this sagittal plane. The reader can easily understand this by, for example, imaging slices of a lemon. The lateral angle indicates the slice in which the sound source is located, and the vertical angle indicates the angle within the slice at which the sound source is located.

In the horizontal plane, however, it is easier to express the direction using the azimuth angle (0°–360°) than to use the combination of the lateral and vertical angles. For example, the expression using the combination of the lateral and vertical angles for the seven directions in the right half of the horizontal plane shown in Fig. 1.5 are (0°, 0°), (30°, 0°), (60°, 0°), (90°, 0°), (60°, 180°), (30°, 180°), and (0°, 180°). These seven directions can be expressed simply using the azimuth angle as (0°, 0°), (30°, 0°), (60°, 0°), (90°, 0°), (120°, 0°), (150°, 0°), and (180°, 0°), respectively. Therefore, in this book, the direction in the horizontal plane is expressed using the azimuth angle.

The azimuth angle ϕ and the elevation angle θ can be transformed into the lateral angle α and the vertical angle β, as follows:

$$\alpha = 90 - \cos^{-1}(\sin\phi\cos\theta)\ (^\circ) \tag{1.3}$$

$$\beta = \sin^{-1}\frac{\sin\theta}{\sqrt{\sin^2\theta + \cos^2\phi\cos^2\theta}}\ (^\circ) \tag{1.4}$$

For instance, the lateral and vertical angles for an azimuth angle of 30° and an elevation angle 45° are 20.7° and 49.1°, respectively.

1.6 Brief History of HRTF Research – Current Achievements and Research Questions to Be Settled

Although the early studies on the HRTF date back to the 1940s, research on HRTFs became active in the 1960s and has progressed by leaps and bounds in the last 60 years. Toward the end of this chapter, I summarize current achievements in research into HRTFs and open questions regarding HRTFs.

1.6.1 Origin of the HRTF

When was the concept of the HRTF proposed? As far as I know, the earliest paper on the HRTF appears to have been written by Wiener and Ross (1946). They measured HRTFs in the horizontal plane in steps of 45° for the frequency range of 200–6000 Hz. However, they did not use the term “HRTF”, but rather “diffraction around the human head”. The first paper to mention the term “HRTF” appears to have been written by Morimoto and Ando (1980).

1.6.2 Physical Characteristics of the HRTF

As mentioned in Sect. 1.1, there exist several distinguishing spectral peaks and notches, which exceed ±10 dB, in the HRTF for the forward direction. For a sound source located in the lateral direction, the spectrum of the HRTF of the ears contralateral to the sound source is flattened, and the level difference between the left and right ears increases. The level difference exceeds 30 dB at particular frequency for a sound source located at an azimuth angle of 90°. The frequencies of spectral notches become higher as a sound source moves from the front to above. However, the frequencies of spectral peaks are approximately constant, regardless of the direction of the sound source. Moreover, there exist significant individual differences in the notch frequency due to individual differences in the shape and size of the pinna (Raykar et al. 2005; Iida et al. 2014).

The notches and peaks are generated in the pinna. At the notch frequency, the anti-nodes are generated at the cymba of concha and the triangular fossa, and a node is generated at the cavity of concha (Takemoto et al. 2012). Peaks are considered to be generated by the resonances of the pinna (Shaw 1997). When the cavities of the pinna are occluded by clay, the notches and peaks vanish (Iida et al. 1998), and the front-back confusion of a sound image increases (Gardner and Gardner 1973; Musicant and Butler 1984; Iida et al. 1998).

1.6.3 Reproduction of the Direction of a Sound Image by Reproduction of the HRTF

Morimoto and Ando (1980) demonstrated that the direction of a sound image can be reproduced by reproducing the listener's own HRTF, using a transaural system (see Sect. 12.2 for details). They also demonstrated that the HRTFs of other listeners often cause front-back confusion of a sound image and inside-of-head localization. Wightman and Kistler (1989a, b) reported that they obtained similar results using headphones with a compensation filter for the transfer function between the headphones and the ear drum.

1.6.4 Cues for the Perception of Lateral Direction

Around 1900, it was known that the interaural time difference and the interaural level difference are cues for the perception of lateral direction (Lord Rayleigh 1877, 1907). Afterwards, around 1960, the quantitative relationships between these interaural differences and the left-right direction of a sound image were reported. Specifically, a sound image is localized at an azimuth angle of 90° for an interaural time difference of 1 ms or an interaural level difference of 10 dB for broad-band noise (Sayers 1964; Toole and Sayers 1965).

1.6.5 Cues for the Perception of Vertical Direction

After the 1970s, a number of studies examined cues for the perception of vertical direction. Consequently, it was found that cues exist in the amplitude spectrum of the HRTF (referred to as spectral cue). More specifically, spectral notches and peaks above 5 kHz were found to play an important role in the perception of vertical direction (Butler and Belendiuk 1977; Musicant and Butler 1984; Hebrank and Wright 1974; Morimoto and Saito 1977; Mehrgardt and Mellert 1977), and the outline of the amplitude spectrum was found to be more important than its fine structure (Asano et al. 1990; Middlebrooks 1992, 1999; Kulkarni and Colburn 1998; Langendijk and Bronkhorst 2002; Macpherson and Sabin 2013). Hebrank and Wright (1974) first mentioned that some particular notches and peaks are more important than others in median plane localization. However, the spectral cues they considered were derived from localization tests using a narrow-band signal, and their hypothesis was found to be incorrect for a broad-band signal (Iida and Ishii 2018). On the other hand, based on sound localization tests using the parametric HRTFs, another hypothesis was proposed, in which the two lowest frequency notches (N1 and N2) above 4 kHz are the important cues (Iida et al. 2007; Iida and Ishii 2018).

Some researchers have reported that the rotation of the head is a cue for front-back discrimination of a sound source (Thurlow and Runge 1967; Perrett and Noble 1997; Kato et al. 2003). For instance, if a sound image moves to the left when you rotate the head to the right, the sound source should be in front of you. However, humans rarely rotate the head spontaneously to identify the direction of a sound source (Nojima et al. 2013). Thus, it is unlikely that humans use head rotation as a cue for front-back discrimination.

1.6.6 Physiological Mechanism for the Perception of the Direction of a Sound Image

It has been reported that there exist functions for calculating the time and level differences between the input signals to the left and right ears, in the medial and lateral nuclei of the superior olive, respectively. The superior olive is located on the pathway from the inner ear to the primary auditory cortex. Furthermore, the dorsal cochlea nucleus (DCN) extracts the spectral notch, and the DCN type IV neurons encode rising spectral edges of the notch (Reiss and Young 2005).

1.6.7 HRTF Models

Research on mathematical and physical models of the HRTF is in progress. For example, the principal component analysis (PCA) model (Kistler and Wightman 1992; Middlebrooks and Green 1992) expresses the amplitude spectrum of an HRTF by superposition of principal components (functions of the frequency domain). Although the PCA model is a beautiful mathematical model, a number of unsolved issues hinder its practical application.

A parametric HRTF model, which focuses on spectral notches and peaks, has also been proposed (Iida et al. 2007). The parametric HRTF is recomposed of some of or all of the spectral notches and peaks extracted from a listener's measured HRTF, regarding the peak around 4 kHz, which is independent of the vertical angle of the sound source (Shaw and Teranishi 1968), as the lower-frequency limit. The notches and peaks are labeled in order of frequency (e.g., P1, N1, P2, N2). The notches and peaks are expressed parametrically in terms of center frequency, level, and sharpness. They carried out sound localization tests in the upper median plane and demonstrated that the parametric HRTF recomposed of only N1, N2, P1, and P2 provided approximately the same localization performance as the measured HRTFs (Iida et al. 2007; Iida and Ishii 2018).

1.6.8 Standardization of the HRTF

There exist remarkable individual differences in HRTFs. This is a serious problem, which prevents acoustic VR from coming into widespread practical use. Current acoustic VR, which can present three-dimensional acoustical sensation to only a specific listener, must be evolved into a universal system that can present three-dimensional acoustical sensation to everyone.

In order to address the individual differences in the HRTFs of different listeners, the standardization of the HRTF has been investigated. A number of dummy-head systems have been developed as hardware solutions. For example, the Knowles Electric Manikin for Acoustic Research (KEMAR) dummy head was developed based on median values measured for dimensions of the head, pinna, and torso of many adults (Burkhard and Sachs 1975). Many of the listeners, however, confuse the front-back direction and/or perceive a sound image inside of the head when they hear sounds using the HRTFs of the KEMAR. This indicates that, at present, the standardization of the HRTF does not solve the problem of individual differences in HRTFs.

1.6.9 Individualization of the HRTF

Another solution to the problem of individual differences in HRTFs is the individualization of the HRTF, which provides personalized HRTFs suitable for each listener. The following approaches have been proposed as individualization methods of the amplitude spectrum of the HRTF:

1. Select the HRTF of the pinna that most closely resembles the pinna shape of the listener from an HRTF database (Zotkin et al. 2003).
2. Expand or compress the amplitude spectrum of the reference HRTF in the frequency domain to minimize the spectral difference between the listener's HRTF and the reference HRTF (Middlebrooks 1999).
3. Synthesize the HRTF using principal components based on the listener's pinna shape (Kistler and Wightman 1992; Middlebrooks and Green 1992).
4. Estimate the notch frequencies of the listener's HRTF from the anthropometry of the listener's pinna, and select the best matching HRTF, the notch frequencies of which are the most similar to the estimated notch frequencies, from an HRTF database (Iida et al. 2014).
5. Select a suitable HRTF from an HRTF database based on a listening test (Seeber and Fastl 2003; Iwaya 2006).

However, problems associated with practical use remain for each method. The hope is that these problems will be resolved in future studies.

1.6.10 Measurement of the HRTF

At this moment, the most reliable method by which to obtain an accurate HRTF is measurement in an anechoic chamber using Eqs. (1.1) or (1.2). However, there are two serious problems in HRTF measurement.

The first problem is related to the need to measure HRTFs with a high SN ratio. In order to accomplish this task, the types of signals suitable for measurement have been investigated. As a result, the maximum-length sequence (MLS) signal and the swept-sine signal were developed. At present, the swept-sine signal has gained mainstream acceptance.

The other problem is related to the need to measure HRTFs for various directions in a short time. A method using reciprocity, which switches the point of the sound source and the point of the receiver, has been proposed (Zotkin et al. 2006). In principle, the HRTFs of many directions can be measured simultaneously by placing a small actuator at the entrance of the ear canal and placing a number of microphones at various directions. However, improvement of the SN ratios of the measured HRTF is still necessary.

1.6.11 Numerical Calculation of the HRTF

In the twenty-first century, numerical calculation of HRTFs by the boundary element method (BEM) and the finite-difference time-domain (FDTD) method became possible using a three-dimensional wireframe model of the head and pinna (Katz 2001; Otani and Ise 2003; Kahana and Nelson 2006; Xiao and Liua 2003; Mokhtari et al. 2007). Using such numerical calculation methods, the understanding of the generation mechanism of the HRTF in the pinna has progressed. However, special equipment, e.g., magnetic resonance imaging (MRI) equipment, is required in order to create a three-dimensional wireframe model of the head and pinnae. Namely, at the moment, calculation of the HRTF is available only to a limited number of listeners. The development of an easy method by which to model the head and pinnae is necessary for the calculation of HRTFs for general listeners.

1.6.12 Directional Band

Blauert (1969/70) reported that there exist particular narrow bands (1/3 octave bands), for which a listener localizes a sound image to a specific vertical angle, regardless of the actual vertical angle of a sound source. He called these bands the directional bands. The center frequencies of the directional bands for the forward and upward directions are 4 kHz and 8 kHz, respectively. The center frequencies for the rearward direction are 1.25 kHz, 10 kHz, and 12.5 kHz. Furthermore, he reported that the band levels of the HRTF of the direction at which the directional bands occurred were higher than the band levels of the HRTFs of other directions. Blauert called these bands the boosted bands. This infers that the spectral peak dominates the perception of vertical direction for a narrow-band signal.

On the other hand, for a wide-band signal, the listener does not localize a sound image to the direction of the directional band, even if the energy of the band corresponding to the directional band is enhanced. For example, when the energy of the band, corresponding to the upward directional band (1/3 octave band of 8 kHz), of a wide-band signal was enhanced and emitted from the forward direction, the listener perceives a single sound image to the front for an enhancement of less than approximately 18 dB. For an enhancement of more than 18 dB, the listener perceives two split sound images. A sound image for the enhanced band is localized above the listener, and another broad-band sound image is localized to the front of the listener.

References

Asano F, Suzuki Y, Sone T (1990) Role of spectral cues in median plane localization. J Acoust Soc Am 88:159–168

Blauert J (1969/70) Sound localization in the median plane. Acust 22: 205–213

Burkhard MD, Sachs RM (1975) Anthropometric manikin for acoustic research. J Acoust Soc Am 58:214–222

Butler A, Belendiuk K (1977) Spectral cues utilized in the localization of sound in the median sagittal plane. J Acoust Soc Am 61:1264–1269

Gardner MB, Gardner RS (1973) Problem of localization in the median plane: effect of pinnae cavity occlusion. J Acoust Soc Am 53:400–408

Hebrank J, Wright D (1974) Spectral cues used in the localization of sound sources on the median plane. J Acoust Soc Am 56:1829–1834

Iida K, Ishii Y (2018) Effects of adding a spectral peak generated by the second pinna resonance to a parametric model of head-related transfer functions on upper median plane sound localization. Appl Acoust 129:239–247

Iida K, Yairi M, Morimoto M (1998) Role of pinna cavities in median plane localization. In: Proceedings of 16th international congress on acoustics. Acoustical Society of America, Seattle, pp 845–846

Iida K, Itoh M, Itagaki A, Morimoto M (2007) Median plane localization using parametric model of the head-related transfer function based on spectral cues. Appl Acoust 68:835–850

Iida K, Ishii Y, Nishioka S (2014) Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae. J Acoust Soc Am 136:317–333

Iwaya Y (2006) Individualization of head–related transfer functions with tournament–style listening test: listening with other's ears. Acoust Sci Tech 27:340–343

Kahana Y, Nelson PA (2006) Numerical modelling of the spatial acoustic response of the human pinna. J Sound Vibration 292:148–178

Kato M, Uematsu H, Kashio M, Hirahara T (2003) The effect of head motion on the accuracy of sound localization. Acoust Sci Tech 24:315–317

Katz BFG (2001) Boundary element method calculation of individual head–related transfer function. I. Rigid model calculation. J Acoust Soc Am 110:2440–2448

Kistler DJ, Wightman FL (1992) A model of head–related transfer functions based on principal components analysis and minimum–phase reconstruction. J Acoust Soc Am 91:1637–1647

Knudsen EI, Konishi M (1979) Mechanisms of sound localization in the barn owl (Tyto alba). J Comp Physiol 133:13–21
Kulkarni A, Colburn HS (1998) Role of spectral detail in sound–source localization. Nature 396:747–749
Kulkarni A, Isabelle SK, Colburn HS (1999) Sensitivity of human subjects to head-related transfer-function phase spectra. J Acoust Soc Am 105:2821–2840
Langendijk EHA, Bronkhorst AW (2002) Contribution of spectral cues to human sound localization. J Acoust Soc Am 112:1583–1596
Lord Rayleigh (1877) Acoustical observations. Phil Mag 3, 6th series: 456–464
Lord Rayleigh (1907) On our perception of sound direction. Phil Mag 13, 6th series: 214–232
Macpherson EA, Sabin AT (2013) Vertical–plane sound localization with distorted spectral cues. Hear Res 306:76–92
Mehrgardt S, Mellert V (1977) Transformation characteristics of the external human ear. J Acoust Soc Am 61:1567–1576
Middlebrooks JC (1992) Narrow–band sound localization related to external ear acoustics. J Acoust Soc Am 92:2607–2624
Middlebrooks JC (1999) Individual differences in external–ear transfer functions reduced by scaling in frequency. J Acoust Soc Am 106:1480–1492
Middlebrooks JC, Green DM (1992) Observations on a principal components analysis of head–related transfer functions. J Acoust Soc Am 92:597–599
Mokhtari P, Takemoto H, Nishimura R, Kato H (2007) Comparison of simulated and measured HRTFs: FDTD simulation using MRI head data. 123rd Audio Engineering Society Convention, New York, Preprint 7240: 1–12
Møller H (1992) Fundamentals of binaural technology. Appl Acoust 36:171–218
Morimoto M, Ando Y (1980) On the simulation of sound localization. J Acoust Soc Jpn (E) 1:167–174
Morimoto M, Saito A (1977) On sound localization in the median plane – effects of frequency range and intensity of stimuli. Technical report of technical committee of psychological and physiological acoustics, Acoust Soc Jpn H-40-1 (in Japanese)
Morimoto M, Joren N, Ando Y, Maekawa Z (1976) On the head-related transfer function. Technical report of technical committee of psychological and physiological acoustics. Acoust Soc Jpn H-31-1 (in Japanese)
Musicant AD, Butler RA (1984) The influence of pinnae-based spectral cues on sound localization. J Acoust Soc Am 75:1195–1200
Nojima R, Morimoto M, Sato H, Sato H (2013) Do spontaneous head movements occur during sound localization? J Acoust Sci & Tech 34:292–295
Otani M, Ise S (2003) A fast calculation method of the head–related transfer functions for multiple source points based on the boundary element method. Acoust Sci Tech 24:259–266
Perrett S, Noble W (1997) The effect of head rotations on vertical plane sound localization. J Acoust Soc Am 104:2325–2332
Raykar VC, Duraiswami R, Yegnanarayana B (2005) Extracting the frequencies of the pinna spectral notches in measured head related impulse responses. J Acoust Soc Am 118:364–374
Reiss LAJ, Young ED (2005) Spectral edge sensitivity in neural circuits of the dorsal cochlear nucleus. J Neuroscience 25:3680–3691
Sayers BM (1964) Acoustic-image lateralization judgement with binaural tones. J Acoust Soc Am 36:923–926
Seeber BU, Fastl H (2003) Subjective selection of non–individual head–related transfer functions. In: Proceedings of the 2003 international conference on auditory display, Boston
Shaw EAG (1997) Acoustical features of the human external ear. In: Gilkey RH, Anderson TR (eds) Binaural and spatial hearing in real and virtual environments. Erlbaum, Mahwah, pp 25–47
Shaw EAG, Teranishi R (1968) Sound pressure generated in an external-ear replica and real human ears by a nearby point source. J Acoust Soc Am 44:240–249

Takemoto H, Mokhtari P, Kato H, Nishimura R, Iida K (2012) Mechanism for generating peaks and notches of head-related transfer functions in the median plane. J Acoust Soc Am 132:3832–3841

Thurlow WR, Runge PS (1967) Effect of induced head movements on localization of direction of sounds. J Acoust Soc Am 42:480–487

Toole FE, Sayers BM (1965) Lateralization judgements and the nature of binaural acoustic images. J Acoust Soc Am 37:319–324

Wiener FM, Ross DA (1946) The pressure distribution in the auditory canal in a progressive sound field. J Acoust Soc Am 18:401–408

Wightman FL, Kistler DJ (1989a) Headphone simulation of free-field listening. I: Stimulus synthesis. J Acoust Soc Am 85:858–867

Wightman FL, Kistler DJ (1989b) Headphone simulation of free-field listening. II: Psychophysical validation. J Acoust Soc Am 85:868–878

Xiao T, Liua QH (2003) Finite difference computation of head–related transfer function for human hearing. J Acoust Soc Am 113:2434–2441

Zotkin DN, Hwang J, Duraiswami R, Davis LS (2003) HRTF Personalization using anthropometric measurements. In: IEEE workshop on applications of signal processing to audio and acoustics

Zotkin DN, Duraiswami R, Grassi E, Gumerov NA (2006) Fast head–related transfer function measurement via reciprocity. J Acoust Soc Am 120:2202–2215