Applied Acoustics 189 (2022) 108610

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Technical note

Toolkit for individualization of head-related transfer functions using parametric notch-peak model

Kazuhiro Iida [a,*], Tsubasa Aizaki [b], Takeshige Kikuchi [b]

[a] Faculty of Advanced Engineering, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino, Chiba 275-0016, Japan
[b] Graduate School, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino, Chiba 275-0016, Japan

ARTICLE INFO

Article history:
Received 25 August 2021
Received in revised form 14 November 2021
Accepted 23 December 2021

Keywords:
Three-dimensional acoustics
Virtual reality
Head-related transfer function
HRTF
Individualization
Toolkit
Notch
Peak

ABSTRACT

Although sound image control and sound field reproduction based on head-related transfer functions (HRTFs) have been studied for years, they have not been put to practical use. The main reason is that individual differences in HRTFs have not been overcome. In the present paper, we propose the parametric notch-peak HRTF model (PNP model), which can continuously change the parameters corresponding to individual differences in HRTFs by using knowledge of spectral notches and peaks and an HRTF database. We then develop a toolkit that can generate individual HRTFs by having the listener adjust the parameters. Finally, localization tests are carried out in order to examine the performance of the individualized HRTFs using the PNP model.

1. Introduction

Accurate three-dimensional sound image localization can be accomplished by reproduction of the listener's own head-related transfer functions (HRTFs) at the entrances of the ear canals [17]. However, there exist remarkable individual differences in HRTFs. The HRTFs of other listeners often cause front-back confusion of a sound image and inside-of-head localization. This is a serious problem, which prevents three-dimensional sound from coming into widespread practical use.

In order to realize three-dimensional sound image localization, it is necessary to provide the listener's own HRTFs, or HRTFs that are suitable for the listener. This process is called individualization of HRTFs.

Methods for obtaining individualized HRTFs for an unknown listener that do not require acoustical measurements can be roughly divided into the following two approaches:

(1) Select a suitable HRTF from an HRTF database.
(2) Generate an individual HRTF from the listener's pinna shape.

In approach (1), the larger the database (the larger the number of HRTFs included), the higher the probability that HRTFs that are suitable for the listener can be selected. However, the time and effort required for the selection process, such as listening tests, increases as the number of HRTFs included in the database increases. In order to solve this problem, a method by which to reduce the total number of listening tests has been studied [21,11].

In order to reduce the size of the database, localization experiments that search for "typical HRTF sets" at the seven directions in the upper median plane were carried out [5]. Typical HRTFs are datasets that include at least one HRTF that provides high sound image localization accuracy for each listener. A total of 34 subjects responded the vertical angle of a sound image for the HRTFs of 100 donors for each of the seven directions. Then, typical HRTF sets were obtained using the results of the localization tests.

Another 34 subjects, who were not involved in the selection process for typical HRTFs, contributed to the validation of typical HRTFs. The mean localization error for the best typical HRTF for each of the 34 subjects decreased as the number of typical HRTFs increased. For n = 10, the mean localization error for the best typical HRTF was less than that for the subject's own HRTF for six of seven target vertical angles, except 0°. For 0°, the mean localization error for the best typical HRTF was 23.8°, whereas the mean localization error was 14.5° for the subject's own HRTF. This implies

* Corresponding author.
E-mail address: kazuhiro.iida@it-chiba.ac.jp (K. Iida).

https://doi.org/10.1016/j.apacoust.2021.108610
0003-682X/

that the selected HRTF from discrete HRTF datasets does not provide as high a localization accuracy as the listener's own HRTF for the front direction.

In approach (2), two methods have been proposed for generation of the amplitude spectra of the individual HRTFs. One is a method that decomposes the amplitude spectrum of the HRTF into several principal components and synthesizes the HRTF using some of the components with weighting coefficients [12,14]. The weighting coefficients depend on both the listener and the direction of a sound source. They have been estimated based on the anthropometry of the listener's pinnae using multiple regression analysis [20,3] or using a deep neural network [4,13]. However, the estimation of the weighting coefficients for an unknown listener has not been successful.

Another method for HRTF individualization estimates the prominent spectral peaks and notches in the individual HRTFs. The minimum HRTF components, which provide approximately the same localization performance as the measured HRTFs, were demonstrated to be the two lowest-frequency notches (N1 and N2) and the two lowest-frequency peaks (P1 and P2) above 4 kHz [6,8].

Each notch and peak (hereinafter referred to as N/P) can be determined by three parameters: center frequency, level, and sharpness (Q factor) by using a peaking filter. Therefore, the generation of individual HRTFs results in the problem of how to set the N/P parameters for each listener and for each direction.

Of these parameters, the frequencies of the N/P for a sound source at the front direction were reported to be estimated based on the anthropometry of the listener's pinnae using multiple regression analysis [7,22,15] or discrimination analysis [9]. However, estimation of the level and Q factor for the individual N/P has not been successful.

In the present paper, we focus on an individual HRTF generation method, in which the listener adjusts the HRTF parameters while listening. However, even the minimum configuration of the HRTF, which is constructed with two notches and two peaks, requires optimization of 12 parameters (three parameters multiplied by four N/Ps), which are not easy to adjust.

In order to solve this problem, we propose a parametric notch-peak HRTF model (PNP model), which reduces the number of independent parameters and which can change the parameters continuously to correspond to individual differences in HRTFs. Furthermore, we have developed a toolkit that generates individualized HRTFs by having the listener adjust the parameters using the PNP model.

2. Generation of individualized HRTFs for the front and rear direction using the PNP model

We consider adaptation of the parameters of an HRTF, which is constructed with two notches and two peaks, to a listener. This is an optimization problem in 12 dimensions (three parameters multiplied by four N/Ps) and is not easy to solve. In order to reduce the number of independent parameters, we tried to estimate some parameters from other parameters, or treat these parameters as constants using the PNP model.

2.1. Outline of the PNP HRTF model

An outline of the procedure of the PNP model is as follows:

(1) Transform the typical HRTFs of 20 ears into parametric HRTFs constructed with two notches (N1, N2) and two peaks (P1, P2) using peaking filters.
(2) Obtain the values of the center frequency, level, and Q factor for each N/P.
(3) Obtain the regression equation for the frequency of each N/P, except for the N2 frequency, with the N2 frequency as an independent variable.
(4) Obtain the mean values of the level and Q factor for each N/P averaged over 20 typical HRTFs.
(5) Generate a modeled HRTF, the center frequency of each N/P of which is obtained by the regression equation with the N2 frequency as an independent variable and the level and Q factor of which are mean values.

In the following section, the details of the PNP model are explained using the front direction as an example.

2.2. Transformation of measured typical HRTFs into parametric HRTFs

The measured typical HRTFs were transformed into parametric HRTFs using peaking filters (second-order IIR filters). Fig. 1 shows an example of the transformation. The green line in Fig. 1(a) shows a measured typical HRTF, and the blue line shows the transformed parametric HRTF. The parameters for the IIR filter (center frequency, level, and Q factor) in Fig. 1(b) were manually set so that the parametric HRTF approximates the measured HRTF.

2.3. Regression analysis of N/P frequency

Fig. 2 shows the distribution of N/P frequencies for the typical HRTFs of 20 ears in the front direction. The N1 and N2 frequencies are distributed between 6,000 and 9,094 Hz (0.60 octaves) and between 9,094 and 11,906 Hz (0.39 octaves), respectively. The P1 and P2 frequencies are distributed between 3,469 and 5,625 Hz (0.70 octaves) and between 7,875 and 10,594 Hz (0.43 octaves), respectively. Such a wide frequency distribution suggests that a typical HRTF may provide accurate sound image localization for each listener.

Regression analysis was performed with the N1 or N2 frequency as an independent variable and the other N/P frequencies as dependent variables. For the front direction, approximately the same correlation coefficients were obtained as those for either the independent variable of N1 or N2, but for the rear target direction, the correlation coefficient obtained as the independent variable for N2 was higher than that for N1. Therefore, in the present study, N2 was used as the independent variable, and other N/P frequencies were estimated using linear regression.

The relationships between the estimated and measured values for the N1, P1, and P2 frequencies are shown in Fig. 3. Although some data points that are far from the diagonal line are observed for P1, most are generally located near the diagonal line. The correlation coefficients for N1, P1, and P2 were 0.80, 0.53, and 0.74, respectively.

Table 1 shows the mean residual errors for N1, P1, and P2. The just-noticeable differences (JNDs) in the N1 frequency for the front direction with regard to vertical localization range from 0.1 to 0.2 octaves for both higher and lower frequencies. The JNDs in the P1 frequency are 0.35 and 0.47 octaves for higher and lower frequencies, respectively [7]. Therefore, the mean residual errors are less than the JNDs.

2.4. Level and Q factor settings for each N/P

For the level and Q factor for each N/P, no high correlation was obtained for other parameters, such as the N2 frequency. However, it has been reported that the sound image could be localized in the front direction if N1 and N2 exceed a certain depth, regardless of the listener's own N1 and N2 depth [16,10]. The Q factor for each

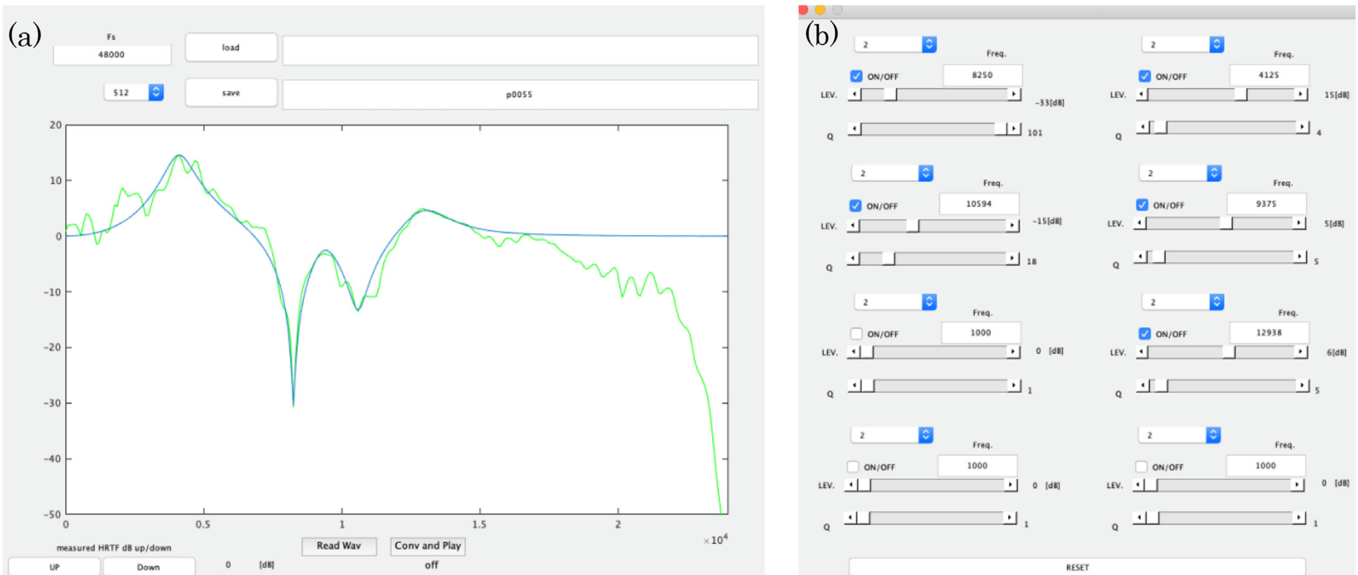

Fig. 1. (a): Examples of measured typical HRTF (green line) and parametrically transformed HRTF (blue line), (b): parameters of IIR filters (center frequency, level, and Q factor). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

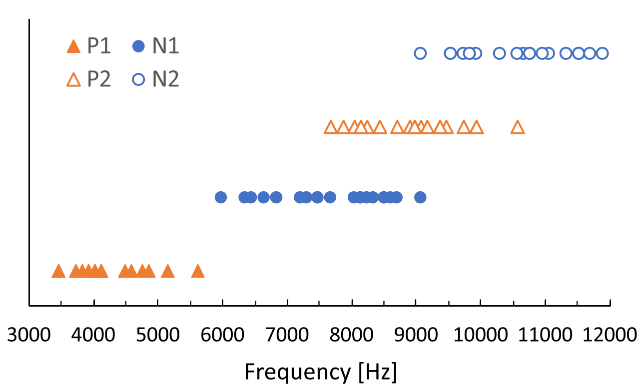

Fig. 2. Distribution of frequencies for N1, N2, P1, and P2 of typical HRTFs for front direction.

N/P was found to be correlated with the level of each N/P. This suggests the possibility of setting a common constant value among listeners.

Based on these findings, the value averaged over 20 typical HRTFs was used as the common constant value among listeners for the level and Q factor for each N/P.

3. Generation of individualized HRTFs for arbitrary direction in horizontal plane

An individualized HRTF for an arbitrary direction in the horizontal plane was generated using the individualized HRTFs for the front and rear directions and the interaural difference cues.

Previous studies have shown that sound image localization in an arbitrary three-dimensional direction can be achieved by adding the interaural difference cues to the spectral cues in the median plane [18]. Based on these findings, the individualized HRTF for an arbitrary direction in the front half of the horizontal plane was generated by adding the interaural time difference (ITD) and the interaural level difference (ILD) to the generated individual HRTF for the front direction, as shown by the blue semicircle in Fig. 4. Similarly, the individualized HRTF for an arbitrary direction in the rear half of the horizontal plane was generated by adding both the ITD and the ILD to the generated individual HRTF for the rear direction, as shown by the green semicircle in Fig. 4.

The ITD was obtained as follows [1]:

$$ITD = (\phi + \sin\phi) \times D/(2c) \quad (1)$$

where ϕ and D denote the incident azimuth angle of a sound in radians and the distance between both ears (diameter of the sphere), respectively.

The ILD varies with both the incident azimuth angle and the frequency of a sound. The results of the experiments on the relationship between the ILD and the lateral localization for a wide-band noise showed that the subjects localized a sound image to the front direction with an ILD of 0 dB and to the lateral direction with an ILD of 10 dB [2]. An approximately linear relationship was observed between the ILD and the azimuth angle of a sound image. In the present study, the ILD was added using this relationship.

4. Development of toolkit for generation of individualized HRTFs

A toolkit, by which a listener individualizes his/her HRTFs while listening to the sound created by the convolution of the sound source and the PNP HRTF model, was developed using MATLAB®.

The user interface for HRTF individualization is shown in Fig. 5. Here, we describe the four steps for individualization of the HRTF for the front direction, as an example.

Step 1: Use the "±500 Hz", "±100 Hz", or "±10 Hz" buttons or the slider to adjust the N2 frequency so that a sound image is localized in the front direction.
Step 2: The "L → R" button provides a stimulus with ITD and ILD added. Use this button to confirm that the sound image moves from the left side to the right side through the front.
Step 3: Adjust the P1 level with the "+1 dB" and "-1 dB" buttons. For the P1 level, the following tendencies were observed in our preliminary experiments. The vertical angle of a sound image rises as the P1 level increases or decreases to the horizontal plane as the P1 level decreases. However, if the P1 level is lowered too much, a sound image is localized close to the head or inside the head.

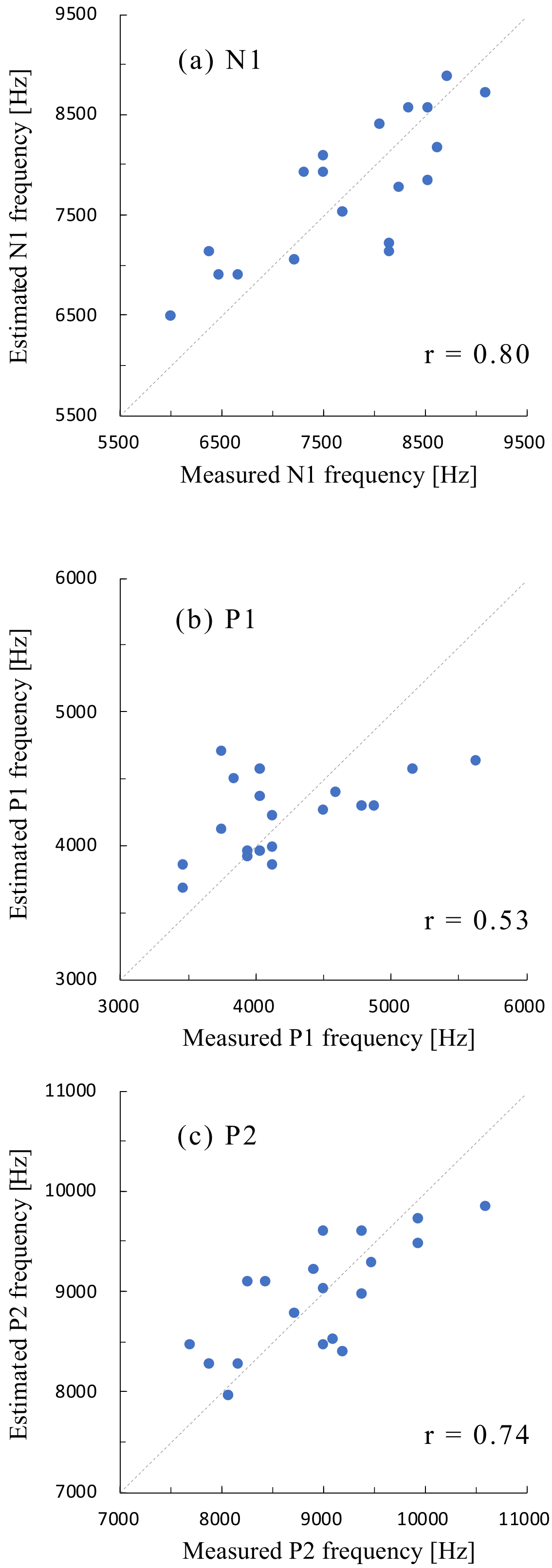

Fig. 3. Relationship between measured and estimated frequency from the N2 frequency. (a) N1, (b) P1, and (c) P2. Here, r denotes the correlation coefficient.

Step 4: The "save" button outputs the individualized HRTF information as head-related impulse responses (HRIRs), together with the parameters of four peaking filters.

Table 1
Mean residual errors for N1, P1, and P2 [octaves].

N1	P1	P2
0.09	0.13	0.07

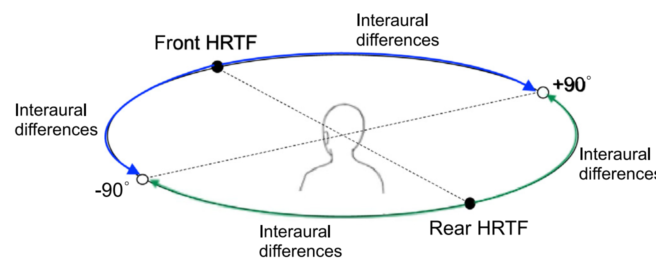

Fig. 4. Sound image control for an arbitrary direction in the horizontal plane using the individualized HRTFs for the front and rear directions and interaural differences.

In addition, after the individualization processing for the front and rear directions, the accuracy of the individualization can be confirmed by the moving sound image function for the 12 directions in the horizontal plane (0°–360°, 30° steps), as shown in Fig. 6.

Each step ends under the following conditions. For Step1, the listener ends the process when a sound image is perceived in the target direction. For Step 2, the listener ends the process when a sound image moves from the left side to the right side through the front. If a sound image moves from the left side to the right side through the inside of the head, go back to Step 1. For Step 3, the listener ends the process when a sound image is localized outside of the head and at the target vertical angle. For the step shown in Fig. 6, the listener ends the process when a sound image moves around the listener in the horizontal plane. If a sound image does not move appropriately, go back to Step 1. Fundamentally, the individualization process starts at Step 1 and ends at Step 4. However, Steps 2 and 3 do not necessarily have to proceed in this order.

5. Examination of performance of HRTF individualization

Localization tests were carried out in order to examine the performance of the individualized HRTFs generated by the toolkit. The individualized HRIRs for each subject were created using the PNP model mentioned in Section 4 in advance of the localization tests. The time required for the tasks in Steps 1 through 4 was approximately 4 min on average for the subjects.

5.1. Method of sound image localization tests

The localization tests were conducted in a quiet soundproof room. The working area of the room was 4.6 m (width) × 5.8 m (depth) × 2.8 m (height). The background A-weighted sound pressure level (SPL) was under 20 dB. Each subject's own HRIRs and the individualized HRIRs using the PNP model were used. The target vertical angles were 0° and 180° in the median plane. The source signal was wideband white noise with a frequency of 200 Hz to 17 kHz. Stimuli were delivered at 63 dB SPL at the entrance of each ear. The duration of the stimuli was 1.2 s, including the rise and fall times, each of which was 0.1 s.

The stimuli, which were obtained by convolution of the sound source and various HRIRs, were presented to the subjects through free air equivalent coupling to the ear (FEC) headphones (beyerdy-

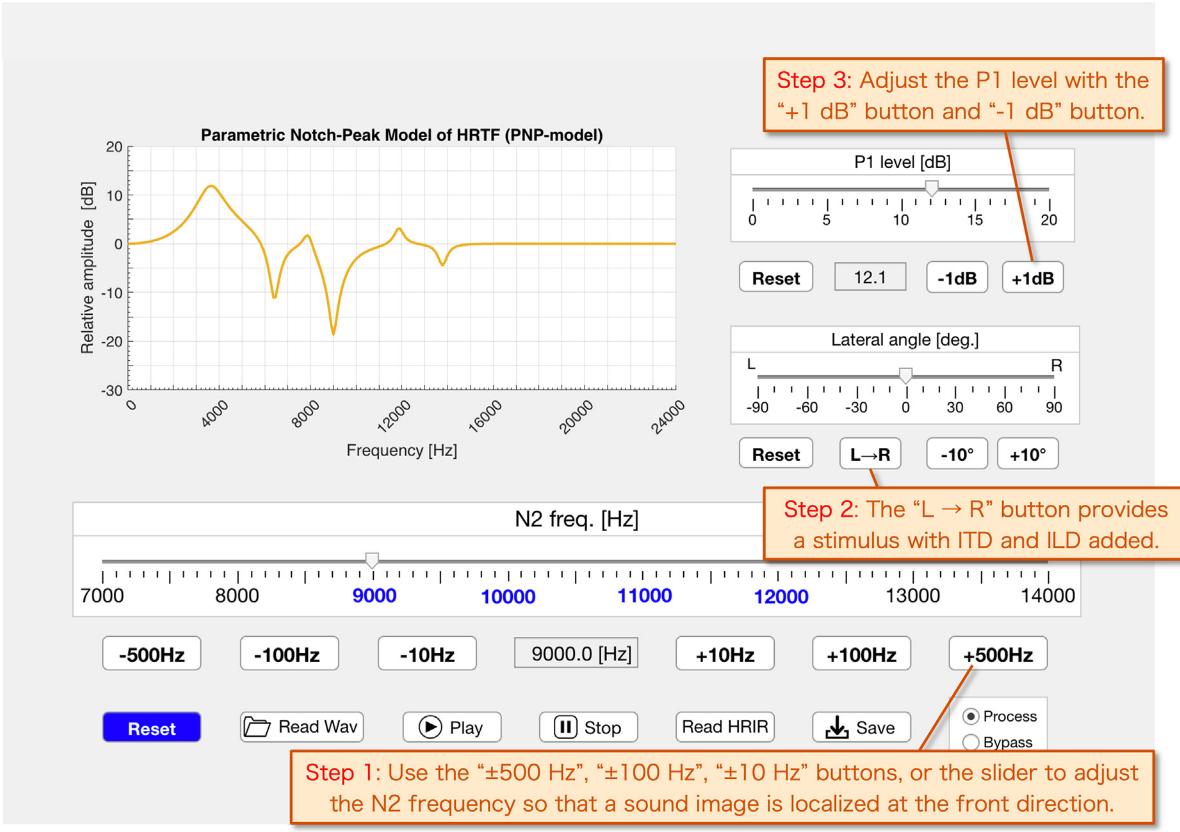

Fig. 5. Graphical user interface of the PNP model for individualization of HRTFs.

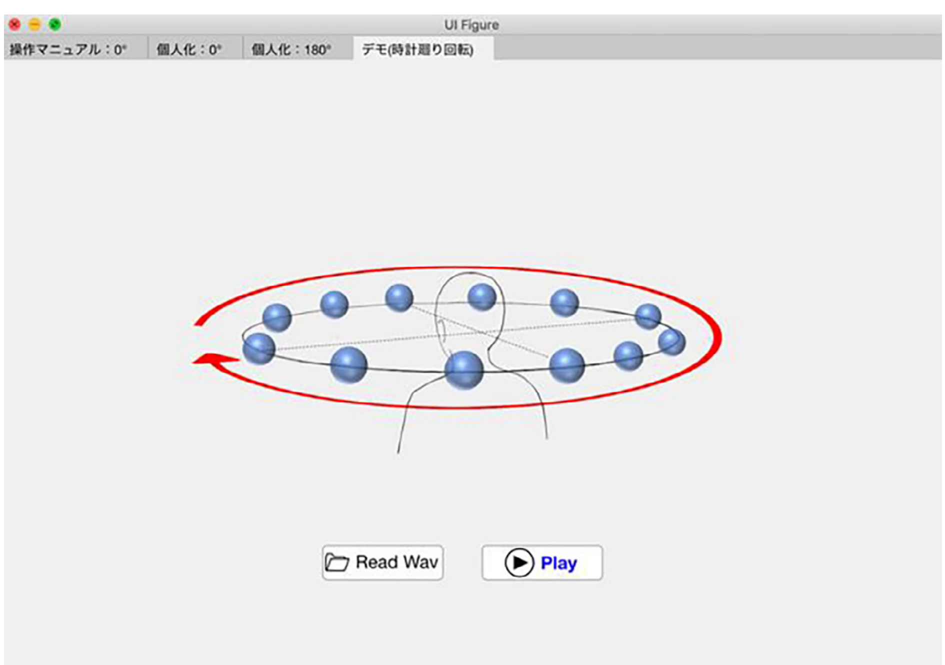

Fig. 6. Snapshot of trial listening of a moving sound image in the horizontal plane for confirmation of HRTF individualization.

namic DT990 PRO) [19]. No compensation of the headphone transfer functions was performed.

The mapping method was adopted as a response method in order to respond on a continuous scale rather than selecting between distinct locations. A circle and a horizontal arrow through its center, which indicated the median plane and the front-back axis, respectively, were displayed on the screen of a laptop computer. The task of the subject was to click on the perceived vertical angle on the circle shown on the computer display using a mouse. Subjects were also instructed to check the box on the display when they perceived a sound image inside their head.

Three subjects participated in the sound localization tests. All of the subjects self-reported normal hearing sensitivity. Each subject responded to each stimulus 10 times in random order. The tests were carried out using a double-blind method.

5.2. Results of sound image localization tests

Fig. 7 shows the responses to the subject's own HRTFs and the individualized HRTFs for each subject. The diameter of each circle is proportional to the number of responses with a resolution of 5°.

For subject A, most of the responses to the subject's own HRTFs and the individualized HRTFs were distributed around the target vertical angles for target vertical angles of both 0° and 180°.

For subject B, most of the responses to the subject's own HRTFs were distributed around the target vertical angles for target vertical angles of both 0° and 180°. The responses to the individualized HRTFs were distributed around the target vertical angle for the target vertical angle of 180°. However, for the target vertical angle of 0°, the responses to the individualized HRTFs shifted slightly upward.

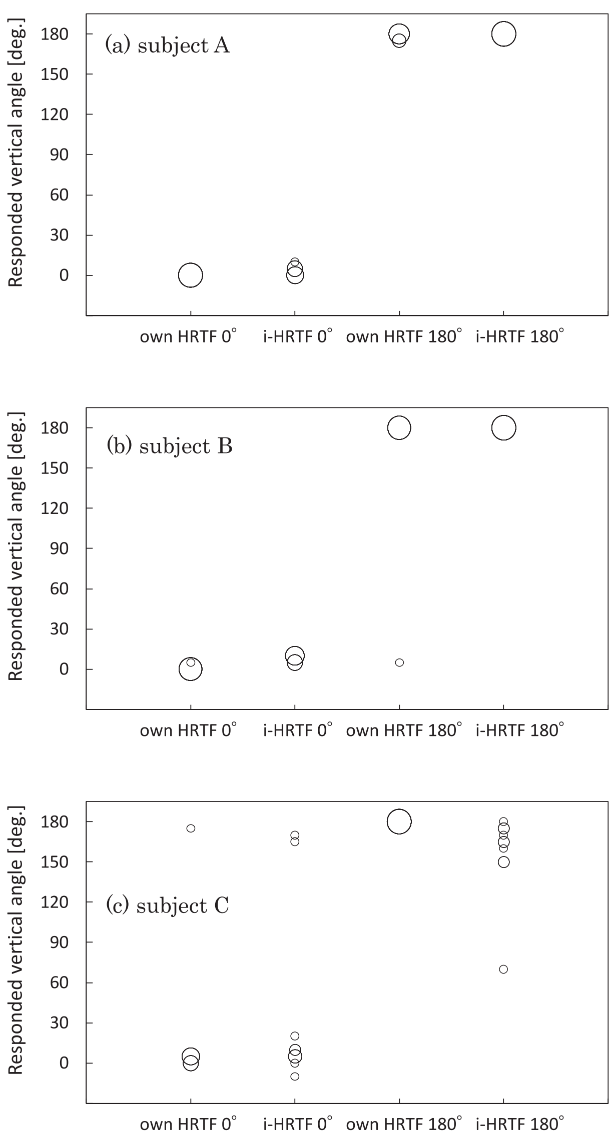

Fig. 7. Responses to the subject's own HRTFs and individualized HRTFs using the PNP model for front and rear directions.

For subject C, most of the responses to the subject's own HRTFs were distributed around the target vertical angles for target vertical angles of both 0° and 180°. The responses to the individualized HRTFs were shifted slightly upward for the target vertical angle of 0°. For the target vertical angle of 180°, the distribution variability of responses to the individualized HRTFs increased.

Table 2 shows the mean elevation error. The mean elevation error is defined as the absolute difference between the responded and target elevation angles averaged over all of the responses. For the front direction, the mean elevation errors for the subject's own HRTFs and individualized HRTFs averaged over three subjects were 1.3° and 6.3°, respectively. For the rear direction, the mean elevation errors for the subject's own HRTFs and individualized HRTFs averaged over three subjects were 1.1° and 7.2°, respectively.

Table 3 shows the ratio of front-back confusion. The ratio of front-back confusion is defined as the ratio of the responses for which the subjects localized a sound image in the quadrant opposite that of the target direction in the upper median plane. For the front direction, the ratio of front-back confusion for the subject's own HRTFs and individualized HRTFs averaged over three subjects were 0.03 and 0.07, respectively. For the rear direction, the ratio of front-back confusion for both the subject's own HRTFs and individualized HRTFs averaged over three subjects was 0.03.

Table 4 shows the ratio of inside-of-head localization. For the front direction, the ratios of inside-of-head localization for the subject's own HRTFs and individualized HRTFs averaged over three subjects were 0.03 and 0.07, respectively. For the rear direction, the ratios of inside-of-head localization for the subject's own HRTFs and individualized HRTFs averaged over three subjects were 0.00 and 0.17, respectively. The ratio of inside-of-head localization of subject C for the individualized HRTF for the rear direction was 0.40. Subject C reported that a sound image was perceived not at the center of the head but at the rear area in the head when he responded as inside-of-head localization.

5.3. Discussions

Tables 5 and 6 show the difference in frequency of notches and peaks between individualized HRTFs and the subject's own HRTFs in the octave scale. Here, the smaller difference for the left and right ears was adopted. The JND of the notch frequency on the vertical angle perception of a sound image is reported to be 0.1–0.2 octaves [7]. Most of the differences were less than the JND. However, the difference in N1 frequency of subject C for the front direction (0.38 octaves) was more than twice the JND. The difference in N2 frequency was also large (0.17 octaves). The individualization process for subject C may have fallen into a locally optimal solution. The mean elevation error, the ratio of front-back confusion, and the ratio of inside-of-head localization of subject C for the front direction were 8°, 0.20, and 0.20, respectively. These values were higher than for other subjects. A possible reason for these values is the difference in N1 and N2 frequencies.

6. Re-examination of performance of HRTF individualization

Individualization was carried out again for subject C. The difference in frequency of notches and peaks between the individualized HRTFs and the subject's own HRTFs is shown in Table 7. The difference for the second individualization decreased compared with the first individualization. For the front direction, the difference in N1 and N2 frequencies were 0.17 and 0.00 octaves, respectively. For

Table 2
Mean elevation error [°].

Subject	Front (0°)		Rear (180°)	
	Own	Individualized	Own	Individualized
A	0.2	3.1	1.8	0.2
B	0.8	7.6	1.2	1.0
C	3.0	8.0	0.3	20.4
Ave.	1.3	6.3	1.1	7.2

Table 3
Ratio of front-back confusion.

Subject	Front (0°)		Rear (180°)	
	Own	Individualized	Own	Individualized
A	0.00	0.00	0.00	0.00
B	0.00	0.00	0.10	0.00
C	0.10	0.20	0.00	0.10
Ave.	0.03	0.07	0.03	0.03

Table 4
Ratio of inside-of-head localization.

Subject	Front (0°)		Rear (180°)	
	Own	Individualized	Own	Individualized
A	0.00	0.00	0.00	0.00
B	0.00	0.00	0.00	0.10
C	0.10	0.20	0.00	0.40
Ave.	0.03	0.07	0.00	0.17

Table 5
Difference in notch frequency and peak frequency between the individualized HRTFs and the subject's own HRTFs for the front direction [octaves].

Subject	N1	N2	P1	P2
A	0.20	0.04	-0.10	0.00
B	0.02	-0.07	-0.10	0.07
C	0.38	0.17	0.13	0.09

Table 6
Difference in frequency of notches and peaks between the individualized HRTFs and the subject's own HRTFs for the rear direction [octaves].

Subject	N1	N2	P1	P2
A	-0.09	-0.11	0.00	-0.01
B	-0.25	-0.14	-0.07	-0.05
C	0.16	0.02	-0.13	0.15

Table 7
Difference in frequency of notches and peaks between the individualized HRTFs (second time) and the own HRTFs for subject C [octaves].

	N1	N2	P1	P2
Front (0°)	0.17	0.00	-0.04	0.03
Rear (180°)	0.03	-0.04	-0.18	0.03

Table 8
Mean elevation error, ratio of front-back confusion, and ratio of inside-of-head localization of the subject's own HRTFs and individualized HRTFs (second time) for subject C.

	Front (0°)		Rear (180°)	
	Own	Individualized	Own	Individualized
Mean elevation error [°]	6.7	19.5	2.8	3.0
Ratio of front-back confusion	0.00	0.00	0.00	0.00
Ratio of inside-of-head localization	0.00	0.00	0.00	0.00

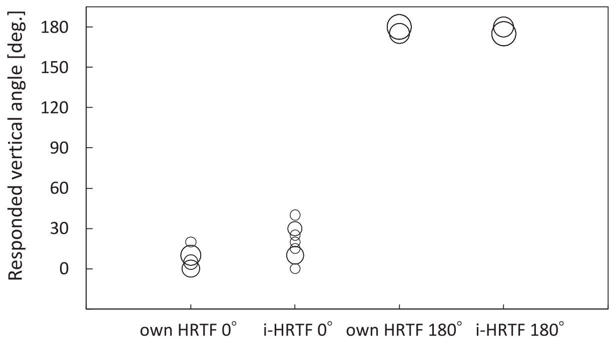

Fig. 8. Responses to the subject's own HRTFs and individualized HRTFs (second time) for subject C using the PNP model for front and rear directions.

the rear direction, the difference in N1 and N2 frequencies were 0.03 and −0.04 octaves, respectively.

Furthermore, localization tests were carried out using the HRTFs generated by the second individualization. The results of the tests are shown in Fig. 8 and Table 8. Neither front-back confusion nor inside-of-head localization were reported for either the front or rear directions. The mean localization error for the front and rear direction was 19.5° and 3.0°, respectively. The error for the front direction was larger compared with the error for the subject's own HRTF (6.7°). For the rear direction, the error was approximately same as the error for the subject's own HRTF (2.8°).

7. Conclusions

In the present paper, we proposed a PNP HRTF model that can continuously change the parameters corresponding to individual differences in HRTFs using knowledge of spectral notches and peaks and an HRTF database. We then developed a toolkit that can generate individual HRTFs by which the listener can adjust the parameters. Finally, localization tests were carried out in order to examine the performance of the individualized HRTFs using the PNP model.

CRediT authorship contribution statement

Kazuhiro Iida: Conceptualization, Methodology, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Tsubasa Aizaki:** Software, Formal analysis, Writing – original draft, Writing – review & editing. **Takeshige Kikuchi:** Software, Formal analysis, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The present study was supported in part by MEXT KAKENHI through a Grant-in-Aid for Scientific Research (C) (Grant Number 19K12068).

References

[1] Blauert J. Spatial hearing revised edition. The MIT Press; 1997. p. 76.
[2] Blauert J. Spatial hearing revised edition. The MIT Press; 1997. p. 158.
[3] Bomhardt R, Braren H, Fels J. Individualization of head-related transfer functions using principal component analysis and anthropometric dimensions. Proc. of meetings on acoustics, 2016.
[4] Chun CJ, Moon JM, Lee GW, Kim NK, Kim HK. Deep neural network based HRTF personalization using anthropometric measurements. Audio Eng Soc Convention 2017;143:9860.
[5] Iida K. In: Head-Related Transfer Function and Acoustic Virtual Reality. Springer; 2019. p. 76–82.
[6] Iida K, Itoh M, Itagaki A, Morimoto M. Median plane localization using parametric model of the head-related transfer function based on spectral cues. Appl Acoust 2007;68:835–50.
[7] Iida K, Ishii Y, Nishioka S. Personalization of head–related transfer functions in the median plane based on the anthropometry of the listener's pinnae. J Acoust Soc Am 2014;136(1):317–33.
[8] Iida K, Ishii Y. 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 2018;129:239–47.
[9] 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. Appl Acoust 2021;177:107929. https://doi.org/10.1016/j.apacoust.2021.107929.
[10] Ishii Y, Gamoh N, Iida K. On a method of individualization of the head-related transfer functions on the basis of the spectral cues and its accuracy, In Proc. Spring Annual Meeting of the Acoust Soc Jpn Mar. 2010;581–584 (in Japanese).
[11] Iwaya Y. Individualization of head–related transfer functions with tournament–style listening test: listening with other's ears. Acoust Sci Technol 2006;27(6):340–3.
[12] Kistler DJ, Wightman FL. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. J Acoust Soc Am 1992;91(3):1637–47.
[13] Miccini R, Spagnol S. HRTF Individualization using Deep Learning. In IEEE 5th VR Workshop on Sonic Interactions in Virtual Environments, Mar. 2020.
[14] Middlebrooks JC, Green DM. Observations on a principal components analysis of head-related transfer functions. J Acoust Soc Am 1992;92(1):597–9.
[15] 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. J Acoust Soc Am 2015;137(2):690–701.
[16] Moore BCJ, Oldfield SR, Dooley GJ. Detection and discrimination of spectral peaks and notches at 1 and 8kHz. J Acoust Soc Am 1989;85(2):820–36.
[17] Morimoto M, Ando Y. On the simulation of sound localization. J Acoust Soc Jpn (E) 1980;1(3):167–74.
[18] Morimoto M, Iida K, Itoh M. Upper hemisphere sound localization using head-related transfer functions in the median plane and interaural differences. Acoust Sci Technol 2003;24:267–75.
[19] Møller H. Fundamentals of binaural technology. Appl Acoust 1992;36(3-4):171–218.
[20] Reddy CS, Hegde RM. A joint sparsity and linear regression based method for customization of median plane HRIR. 49th Asilomar conference on signals, systems, and computers. Nov. 2015;785-789.
[21] Seeber BU, Fastl H. Subjective selection of non–individual head–related transfer functions. In: Proceedings of the 2003 international conference on auditory display, Boston, 2003.
[22] Spagnol S, Avanzini F. Frequency estimation of the first pinna notch in head-related transfer functions with a linear anthropometric model. In Proc. the 18th Int. conference on digital audio effects (DAFx-15), 2015. Trondheim, Norway.