

# 耳介形状による受聴者個人の正中面頭部伝達関数生成の試み

飯田 一博<sup>†</sup> 島崎 光<sup>‡</sup> 太田 正人<sup>‡</sup>

<sup>†</sup> 千葉工業大学工学部 〒275-0016 千葉県習志野市津田沼 2-17-1

<sup>‡</sup> 千葉工業大学大学院工学研究科

E-mail: <sup>†</sup>kazuhiro.iida@it-chiba.ac.jp

あらまし 3次元音響再生や音のVRの実現をめざして, 頭部伝達関数(HRTF)の個人化の研究が国際的に進められている。本研究では, 上半球正中面の初期HRTFの各離散周波数における振幅レベルを目的変数, 受聴者の耳介形状パラメータを説明変数とした重回帰分析により, 個人のHRTFを生成することを試みた。その結果, 耳介形状から生成したHRTFには, 実測初期HRTFにある特徴的なノッチやピークが観察された。これにより, HRTFデータベースなしでの個人のHRTF生成の可能性の一端が示された。

キーワード 頭部伝達関数, 生成, 個人化, 耳介形状, 正中面

## An approach to generation of the individual head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae

Kazuhiro IIDA<sup>†</sup>, Hikaru SHIMAZAKI<sup>‡</sup>, and Masato OOTA<sup>‡</sup>

<sup>†</sup> Faculty of Engineering, Chiba Institute of Technology 2-17-1 Tsudanuma, Narashino, Chiba, 275-0016 Japan

<sup>‡</sup> Graduate School, Chiba Institute of Technology

E-mail: <sup>†</sup>kazuhiro.iida@it-chiba.ac.jp

**Abstract** In order to address the individual differences in the head-related transfer functions (HRTFs) of different listeners, the individualization of the HRTF has been investigated. In the present study, multiple regression analyses were carried out as objective variables of the amplitude level of each discrete frequency of the early HRTFs in the upper median plane and as explanatory variables of fourteen anthropometric parameters of the pinnae. The results showed the potential of generation of the individual HRTFs from the listener's anthropometry of the pinnae without any HRTF database.

**Keywords** Head-Related Transfer Function, Generation, Individualization, Pinna, Anthropometry, The Median Plane

### 1. はじめに

頭部伝達関数(HRTF: head-related transfer function)には顕著な個人差が存在する[1]。本人のHRTFを再現すると受聴者は目標方向に音像を知覚するが, 他人のHRTFを再現すると頻繁に音像の前後誤判定や頭内定位が生じる[2]。これは3次元音響再生や音のバーチャルリアリティ(VR)の実用化を阻害する主たる要因の1つである。不特定多数の受聴者に対して音のVRを実現するにはHRTFの個人差の克服は避けて通れない。

受聴者に適合するHRTFを提供するために様々な研究が進められているが, 以下の2つに大別できる。

1) できるだけ多くの受聴者に適合するできるだけ少数のロバストなHRTFを選出する。

2) 各受聴者に適合する個別のHRTFを選出もしくは生成する。

1) において, Møller *et al.* [3]は20名の被験者に対して, 本人を含む30名の提供者のHRTFを用いて

ヘッドホン再生により音像定位実験を行い, 最も音像定位誤差が小さくなる提供者を選出し, Typical subjectと呼んだ。ランダムな提供者のHRTFの前後誤判定率の平均値が36.3%であるのに対し, Typical subjectのHRTFに対する前後誤判定率は21.2%であった。さらに, われわれは上半球正中面において, 100名のHRTF提供者と74名の被験者を用いて, 不特定多数の受聴者に精度の高い音像定位を実現するのに必要な複数(6-10種類程度)のロバストなHRTF(Typical HRTF)を選出した[4]。

一方2)については, 筆者が知る限り以下のような方法が提案されてきた。

① 受聴者の耳介形状に近い耳のHRTFを用いる[5]。

② 標準的なHRTFを受聴者の耳介形状に応じて周波数軸上で伸縮(scaling)する[6]。

③ 受聴者の耳介形状を用いてPCA(Principle Component Analysis)によりHRTFを合成する[7,8]。

④ 受聴者の耳介形状からスペクトラルキューを推定

し、それに近い HRTF をデータベースから選出する [9].

⑤ 試聴によりデータベースから選出する [10,11].

これらのうち、実用化をめざして現在も継続して研究が進められているのは③と④である。③は HRTF を複数の周波数の関数に分解し、そのうちの主要な関数で再構成する際に、音源方向と個人の耳介形状に由来する重み係数を掛ける [8]. しかし、この重み係数の推定に課題が残されている。

一方④では、まず HRTF を複数のノッチとピークに分解して、必要最小限のノッチとピークの構成 (N1, N2, P1, P2) を明らかにした [12,13]. ノッチやピークの周波数については、概ね弁別閾内 (0.1-0.2 オクターブ) の誤差で耳介形状から推定可能になってきたが [9,14], ノッチやピークのレベルの推定は解決すべき課題として残されている。

本研究では、これら従来の方法とは異なる、新たな個人化 HRTF の生成方法を提案し、その可能性を検証した。

## 2. 提案法の考え方

提案法 (GAP: Generation of the Amplitude spectrum from the Pinna anthropometry) では、ノッチやピークに限定せず、可聴域の各離散周波数において、HRTF の振幅レベルを受聴者の耳介形状から推定し、個人の HRTF の振幅スペクトルを生成する。

Fig.1 に HRTF データベースからランダムに選んだ被験者 10 人の HRTF (正面方向) の振幅スペクトルを示す。4 kHz 以上で顕著な個人差が観察される。この帯域では、波長との関係から耳介形状の個人差が強く影響していると考えられる。

ここで、離散化した (例えば 48000Hz/512 = 93.75 Hz 間隔) 各周波数において、HRTF の振幅レベルを耳介形状で説明することを考える。Fig.1 の各周波数において、縦軸に沿った音圧レベルの分布を受聴者の耳介形状から推定する。このような推定モデル (例えば重回帰モデル) を一旦構築すれば、HRTF データベースがなくても、偏回帰係数と受聴者の耳介形状データだけで個人の HRTF を生成することが可能になる。

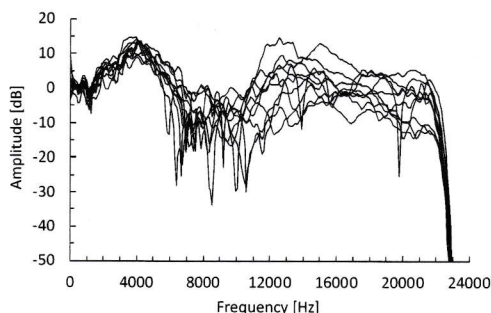


Fig.1 Individual difference of the amplitude spectrum of the HRTF for the front direction among ten subjects.

## 3. HRTF の前処理

HRTF のノッチやピークは耳介で生成されることが知られている [15,16]. また、耳介の影響は HRTF の観測点 (外耳道入口) との位置関係から、主に頭部インパルス応答 (HRIR: head-related impulse response) の初期部分に含まれる。したがって、HRIR の初期部の振幅スペクトルは耳介形状と相関が高いと考えられる。さらに、振幅の絶対値が最大となるサンプルの  $\pm 1$ ms の範囲を切り出した HRIR でも通常の HRIR と同等の音像定位精度が得られることが報告されている [17].

そこで、無響室で測定した成人男女 27 名 (54 耳) の上半球正中面 7 方向 (上昇角:  $0^{\circ}$ - $180^{\circ}$ ,  $30^{\circ}$  間隔) の HRIR に以下の前処理を施した [17].

- 1) HRIR の振幅の絶対値が最大となるサンプル ( $S_{\max}$ ) を検出する。
- 2)  $S_{\max}$  を中心として 4 項、96 (時間軸上で前に 48, 後ろに 48) サンプルの Blackman-Harris 窓で初期 HRIR を切り出す (fs: 48 kHz).
- 3) 512 サンプルの FFT により振幅スペクトルを算出する (周波数分解能: 93.75 Hz).

初期 HRTF の算出例を Fig.2 に示す。初期 HRTF (破線) では微細なスペクトル変動が消滅し、主要なノッチやピークの概形が保持されていることが観察される。

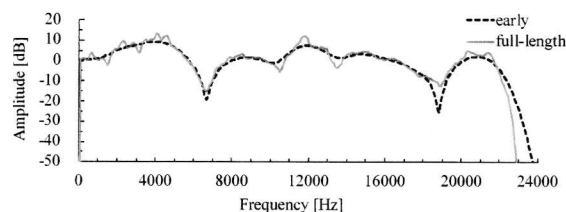


Fig.2 Example of the amplitude spectrum of the early HRTF (broken line) and the full-length (usual) HRTF (solid line).

## 4. 耳介形状パラメータ

初期 HRTF を測定した 54 耳について、14 箇所の耳介形状パラメータ ( $x_1$ - $x_{14}$ ) を求めた。従来の多くの研究で用いられている耳介形状パラメータ [5,9] には次の 2 つの問題があった。1) 外耳道入口と各部位との位置関係を表してない。2) 各部位の計測が容易ではない。

本研究では、Fig.3 に示すように、まず外耳道入口を原点 ( $p_0$ ) として  $p_1$ - $p_{12}$  の 2 次元座標を求めた。 $p_1$ - $p_{12}$  は  $C_1$  (耳輪の内側境界線),  $C_2$  (対輪),  $C_3$  (耳甲介の外側境界線) と、原点を通る上昇角  $120^{\circ}$ - $270^{\circ}$  ( $30^{\circ}$  間隔) の直線との交点である。 $p_0$  から  $p_1$ - $p_{12}$  までの距離を  $x_1$ - $x_{12}$  とした。 $x_{13}$  は耳介の傾き、 $x_{14}$  は耳甲介腔の深さである。 $x_1$ - $x_{13}$  は被験者の耳介写真および横顔の写真から計測し、 $x_{14}$  はノギスで計測した。計測結果を Table 1 に示す。 $x_{13}$  を除き Max/Min は 1.5-2.6 であった。 $x_{13}$  の個人差は大きく、Max/Min は 10.0 であった。

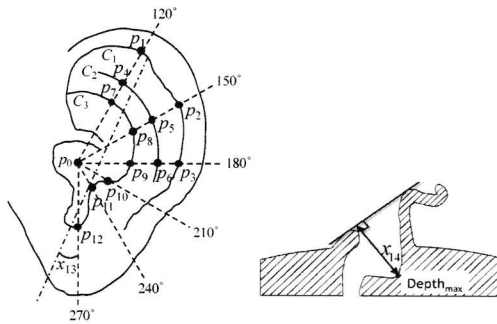


Fig.3 14 anthropometric parameters of the pinna.

Table 1 Pinnae dimensions of 54 ears [mm]

	X1	X2	X3	X4	X5	X6	X7
Ave.	35.1	29.0	22.9	25.3	21.9	18.2	15.9
Max.	44.6	37.2	28.6	31.4	26.8	22.9	20.3
Min.	26.4	21.8	17.6	18.8	18.4	14.1	11.0
Max/Min	1.7	1.7	1.6	1.7	1.5	1.6	1.8

	X8	X9	X10	X11	X12	X13[°]	X14
Ave.	18.5	15.3	9.5	9.0	10.9	22.4	13.8
Max.	21.9	19.1	14.1	12.9	13.1	40.0	17.6
Min.	14.5	9.7	5.3	5.3	7.8	4.0	9.7
Max/Min	1.5	2.0	2.6	2.4	1.7	10.0	1.8

### 5. 重回帰分析による HRTF 推定

初期 HRTF の振幅レベルを目的変数, 14 種類の耳介形状パラメータを説明変数として, 各上昇角の各離散周波数で重回帰分析を行った(式(1)).

$$y(s, \beta, f) = \sum_{i=1}^{14} a_i(\beta, f)x_i(s) + b(\beta, f) \quad (1)$$

ここで,  $s, \beta, f, a_i, x_i, b$  は被験者, 上昇角 ( $0^\circ$ - $180^\circ$ ,  $30^\circ$ 間隔), 離散周波数, 偏回帰係数, 耳介形状パラメータ, 定数を表す. 総当たり法を用いて, 自由度調整済み決定係数が最大となる耳介形状パラメータを選択した.

Fig.4 に正面方向の振幅レベルの実測値と推定値の関係の例を示す. (a)3kHz は第 1 ピーク (P1) 付近であり, 個人差は小さい. (b)12kHz では被験者によりノッチの場合もあればピークの場合もあり, 個人差が大きい. 重相関係数はそれぞれ 0.57 および 0.60 であった.

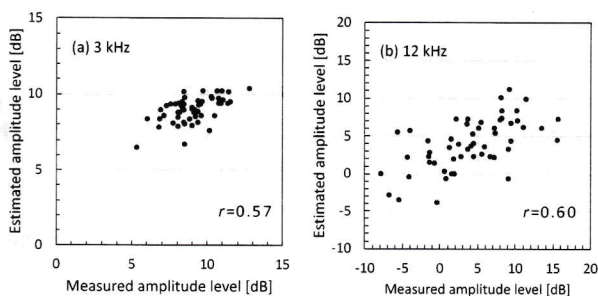


Fig.4 Relationship between the measured amplitude level and the estimated amplitude level at 3 kHz and 12 kHz.

重相関係数を周波数の関数として表した例を Fig.5 に示す(上昇角  $0^\circ$ ). 周波数により差異はあるが, 重相関係数は概ね 0.3-0.7 であった.

各上昇角の 20 kHz 以下の重相関係数の平均値は 0.53-0.61 となり (Table 2), 上方で高く, 前方や後方でやや低くなる傾向がみられた. これは, 上方と比べて, 前方や後方のスペクトルの起伏が激しいことに依ると考えられる.

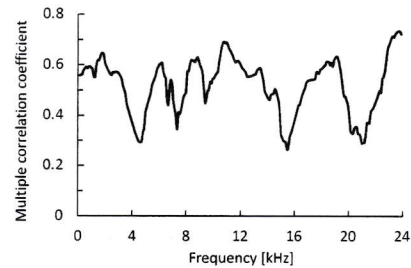


Fig.5 An example of multiple correlation coefficient as a function of frequency for vertical angle of  $0^\circ$ .

Table 2 Multiple correlation coefficient averaged over 93.75 Hz to 19,968.75 Hz.

	$0^\circ$	$30^\circ$	$60^\circ$	$90^\circ$	$120^\circ$	$150^\circ$	$180^\circ$
	0.53	0.55	0.61	0.59	0.56	0.55	0.55

重回帰分析により推定した振幅スペクトルの例(2耳)を Fig.6 に示す. 初期 HRTF の実測値(実線)にみられる特徴的なノッチやピークが推定値(破線)でも観察される. 詳しくみると, 第 1 ピーク (P1) については, 周波数, レベルともに精度よく推定できているが, P2 のレベルには推定誤差がみられる. またノッチについては, N1, N2 ともに周波数は精度よく推定されているが, レベルには推定誤差がみられる.

さらに, スペクトルの推定精度を定量的に検証するため SD (spectral distortion, 式(2))を算出した. ここで  $HRTF_e, HRTF_m, f$  はそれぞれ推定 HRTF, 実測初期 HRTF, 離散周波数を示す.

各上昇角における 54 耳の SD の平均値は 3.2-4.2dB であった (Table 3). スペクトルが平坦な上方では低く, ノッチやピークが顕著な前方や後方では高くなる傾向がみられた.

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N \left[ 20 \log_{10} \left| \frac{HRTF_e(f_i)}{HRTF_m(f_i)} \right| \right]^2} \quad (2)$$

Table 3 SD between the measured early HRTF and the generated HRTF averaged over 54 ears (93.75 Hz to 19,968.75 Hz), in decibel.

	$0^\circ$	$30^\circ$	$60^\circ$	$90^\circ$	$120^\circ$	$150^\circ$	$180^\circ$
	4.2	3.7	3.6	3.2	4.0	3.7	3.8

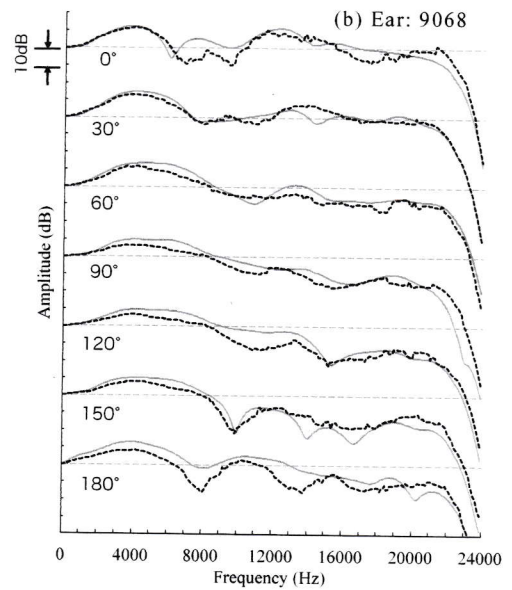
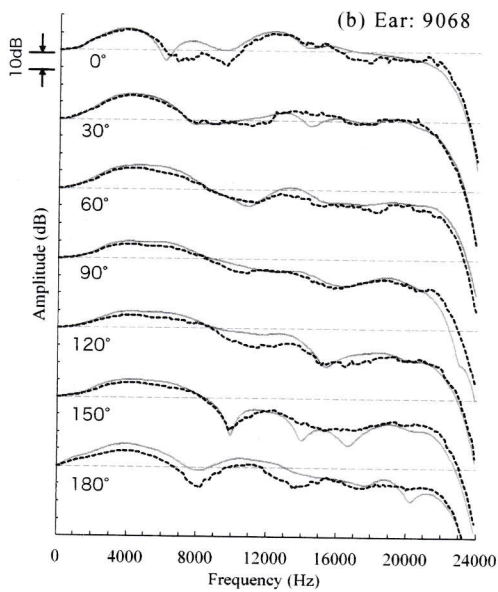
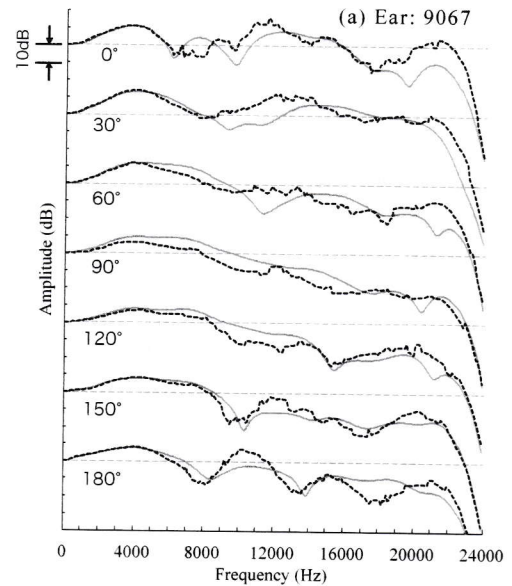
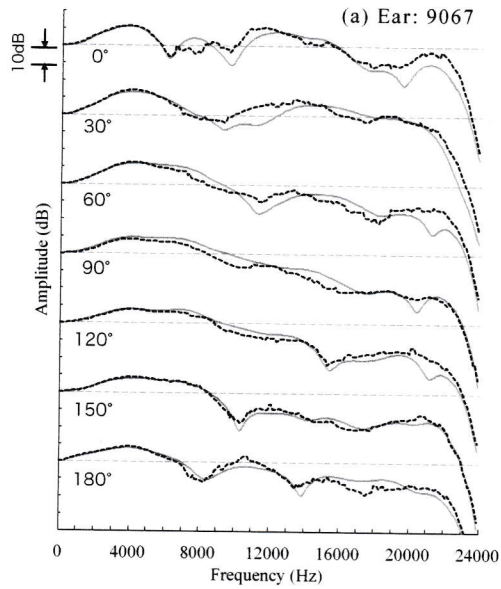


Fig.6 Examples of the amplitude spectrum of the measured early HRTFs (solid line) and the estimated HRTFs (broken line) in the upper median plane.

Fig.7 Examples of the amplitude spectrum of the measured early HRTFs (solid line) and the generated HRTFs (broken line) of a naïve subject.

## 6. ナイーブな被験者の HRTF 生成

重回帰分析に用いていないナイーブな被験者に対する HRTF 生成精度を検証した。27 名(54 耳)の被験者のうち、24 名(48 耳)の初期 HRTF を重回帰分析に用い、残りの 3 名(6 耳)をナイーブな被験者として HRTF を生成した。ナイーブな被験者の HRTF は、その耳介形状パラメータと、48 耳の重回帰分析で得た偏回帰係数を用いて生成した。

生成したナイーブな被験者の HRTF の一例(Fig.6 の重回帰分析と同じ 2 耳)を Fig.7 に示す。Fig.6 の推定値と比較すると精度はやや低下したが、特徴的なノッチやピークは再現されている。

各上昇角における 6 耳の生成 HRTF と実測初期 HRTF の SD の平均値は 4.2-6.4dB であった(Table 4)。Table 3 と同様にスペクトルが平坦な上方では低く、ノッチやピークが顕著な前方や後方では高くなる傾向がみられた。また、SD が大きな値となる耳は、その耳介形状パラメータが中央値から離れている傾向があった。今後、さらに詳細な原因の分析と対策を進めたい。

Table 4 SD between the measured early HRTF and the generated HRTF averaged over 3 naïve subjects (6 ears), in decibel.

0°	30°	60°	90°	120°	150°	180°
6.4	5.5	4.8	4.2	5.2	5.2	6.4

## 7. 結論

正中面の初期 HRTF の各離散周波数における振幅レベルを受聴者の耳介形状から推定することにより、HRTF データベースなしで個人の HRTF の振幅スペクトルを生成することを試みた。その結果、以下の知見が得られた。

1) 耳介形状から重回帰分析により推定した HRTF には、実測初期 HRTF にある特徴的なノッチやピークが観察された。ピーク・ノッチの周波数については概ね精度よく推定されたが、ノッチのレベルについては推定誤差が観察された。

2) 重回帰分析に用いた被験者の推定 HRTF と実測初期 HRTF の平均 SD は正中面 7 方向で 3.2-4.2dB であった。

3) 耳介形状パラメータから生成したナイーブな被験者の HRTF の精度は 1) の推定精度と比較するとやや低下したが、特徴的なノッチやピークは再現された。

4) ナイーブな被験者の生成 HRTF と実測初期 HRTF の平均 SD は正中面 7 方向で 4.2-6.4dB であった。

以上より、HRTF データベースなしでの個人の HRTF 生成の可能性の一端が示された。

## 8. 今後の展開

今後、さらに以下の取り組みが必要である。

- ① 物理的な HRTF 生成精度の向上。特に生成誤差の大きい耳の原因分析と対策。
- ② NFD (Notch Frequency Distance)[9] や NLD (Notch Level Distance) など上昇角知覚の手掛かりに着目した物理指標による生成精度の評価。
- ③ 音像定位実験による生成 HRTF の知覚面での精度の検証。
- ④ 耳介形状データの簡易計測方法の確立。

## 文 献

- [1] 飯田一博, 頭部伝達関数の基礎と 3 次元音響再生への応用, pp.63-70, コロナ社, 東京, 2017.
- [2] M. Morimoto and Y. Ando, "On the simulation of sound localization," J. Acoust. Soc. Jpn. (E) vol.1, pp.167-174, 1980.
- [3] H. Møller, C. B. Jensen, D. Hanmmershøi, and M. F. Sørensen, "Using a typical human subject for binaural recording," Audio Eng. Soc., Reprint 4157(C-10), 1996.5.
- [4] 吉井菜実, 柴田真由, 飯田一博, "上半球正中面における Typical 頭部伝達関数の選出," 音講論(秋), 1-6-5, 2018.9.
- [5] D. N. Zotkin, J. Hwang, R. Duraiswami, and L. S. Davis, "HRTF Personalization using anthropometric measurements," IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2003.
- [6] J. C. Middlebrooks, "Individual differences in external-ear transfer functions reduced by scaling in frequency," J. Acoust. Soc. Am., vol.106, pp.1480-1492, 1999.
- [7] D. J. Kistler and F. L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," J. Acoust. Soc. Am., vol.91, pp.1637-1647, 1992.
- [8] R. Bomhardt, M. Klein, and J. Fels, "A high-resolution head-related transfer function and three-dimensional ear model database," Proc. Meetings on Acoustics, 29, 050007, 2017.
- [9] K. Iida, Y. Ishii, and S. Nishioka, "Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae," J. Acoust. Soc. Am., vol.136, pp.317-333, 2014.
- [10] B. U. Seeber and H. Fastl, "Subjective selection of non-individual head-related transfer functions, Proc. of the 2003 International Conference on Auditory Display, Boston, MA, USA, 2003.
- [11] Y. Iwaya, "Individualization of head-related transfer functions with tournament-style listening test: Listening with other's ears, Acoust. Sci. & Tech., vol.27, pp.340-343, 2006.
- [12] K. Iida, M. Itoh, A. Itagaki, and M. Morimoto, "Median plane localization using parametric model of the head-related transfer function based on spectral cues," Appl. Acoust. vol.68, pp.835-850, 2007.
- [13] K. Iida and Y. Ishii, "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., vol.129, pp.239-247, 2018.
- [14] 島崎光, 飯田一博, "受聴者の耳介形状による上半球正中面の初期頭部伝達関数のノッチ周波数の推定," 音講論(春), pp.1585-1588, 2018.
- [15] E. A. G. Shaw, "Acoustical features of the human external ear," Binaural and spatial hearing in real and virtual environments, ed. R. H. Gilkey and T. R. Anderson, Erlbaum, Mahwah, New Jersey, pp.25-47, 1997.
- [16] H. Takemoto, P. Mokhtari, H. Kato, R. Nishimura, and K. Iida, "Mechanism for generating peaks and notches of head-related transfer functions in the median plane," J. Acoust. Soc. Am., vol.132, pp.3832-3841, 2012.
- [17] K. Iida and M. Oota, "Median plane sound localization using early head-related impulse response," Appl. Acoust., vol.139, pp.14-23, 2018.