

受聴者の耳介形状による上半球正中面の頭部伝達関数生成の試み*

☆島崎光, 太田正人, 飯田一博 (千葉工大)

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発生する[2]. これは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を提供する方法は活発に研究されているが, 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
- 2) については, 筆者が知る限りにおいて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を用いる方法
 - 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)する方法
 - ③受聴者の耳介形状を用いてPCA(Principle Component Analysis)によりHRTFを合成する方法
 - 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を選出する方法
 - ⑤試聴によりHRTFを選出する方法

これらのうち, 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という考え方である[3]. しかし, この重み係数の推定に課題が残されている.

④では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(N1, N2,

P1, P2) [5,6] で再構成し, これら4つのファクタを耳介形状から推定しようというものである[4,7]. しかし, ノッチやピークの推定は部分的な成功に留まっている. ノッチやピークの中心周波数については, 概ね弁別閾内の誤差で推定可能になってきたが, 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, 上記従来の方法とは異なる個人の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

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ことを考える.

Fig.1にランダムに選んだ被験者10人の左耳の正面方向のHRTFを示す. 4kHz以上の帯域では個人差が顕著である. 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, 離散化した各周波数(例えば $48000/512=93.75$ Hz 間隔)におけるHRTFの振幅レベルの分布を耳介形状で説明することを試みる. 各音源方向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で説明する重回帰モデルを構成できれば, 耳介形状データのみによる個人のHRTF生成の可能性が生まれる.

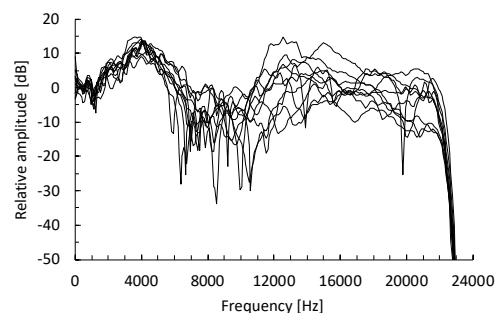


Fig.1 Individual difference of HRTF for the front direction among ten subjects.

* An approach to generate the head-related transfer functions in the upper median plane based on the anthropometry of the listener's pinnae, by SHIMAZAKI, Hikaru, OOTA, Masato, and IIDA, Kazuhiro (Chiba Institute of Technology)

3 HRTF の前処理

ノッチやピークは耳介で生成される[8,9].
 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、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。また、HRIR の継続長をパラメータとした音像定位実験により、振幅の絶対値が最大をとるサンプルの±1ms の応答だけで通常の 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[10].

そこで本研究では、無響室で成人男女 27 名(54 耳)の上半球正中面 7 方向(上昇角:0°-180°, 30°間隔)の HRTF を測定し、以下に示す前処理を施した[9].

- 1) 頭部インパルス応答の振幅の絶対値が最大となるサンプル(Smax)を検出する。
- 2) Smax 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 により初期 HRTF の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。

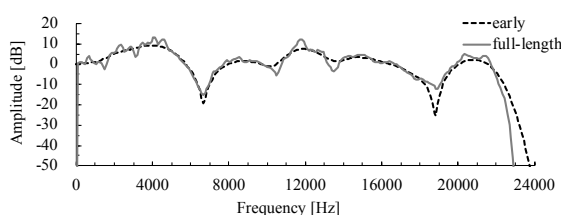


Fig.2 Example of amplitude spectrum of early HRTF (broken line) and Full-length (usual) HRTF (solid line).

4 耳介形状パラメータ

初期 HRTF を求めた 54 耳について Fig.3 および Table 1 に示す 13 箇所の耳介形状パラメータを計測した。このうち、 x_1 - x_9 は被験者の耳型よりノギスで計測し、 x_{10} - x_{13} は被験者の横顔および耳介の写真から計測した。計測結果を Table 2 に示す。 x_1 - x_9 および x_{11} - x_{13} の個人差は 3.3-25.0 mm, x_{10} の個人差は 36°であった。

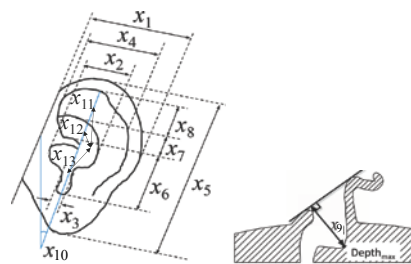


Fig.3 13 anthropometric parameters of the pinna.

Table 1 Definition of 13 anthropometric parameters of the pinna.

計測箇所	名称
x_1	最大耳幅
x_2	耳甲介腔の最大幅
x_3	珠間切痕の最大幅
x_4	耳輪の最大幅
x_5	最大耳長
x_6	耳甲介腔の長さ
x_7	耳甲介舟の長さ
x_8	舟状窩の高さ
x_9	耳甲介腔の深さ
x_{10}	耳介の傾き
x_{11}	外耳道入口から三角窩までの長さ
x_{12}	外耳道入口から耳甲介舟までの長さ
x_{13}	外耳道入口から耳甲介腔までの長さ

Table 2 Range of pinnae dimensions for 54 ears [mm]

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
Min.	31.2	14.3	5.5	18.1	58.2	17.7	2.6
Max.	43.8	21.8	11.5	31.8	83.2	25.1	9.9
Range	12.6	7.5	6.0	13.7	25.0	7.4	7.3

	x_8	x_9	$x_{10}[^\circ]$	x_{11}	x_{12}	x_{13}
Min.	13.2	9.7	4.0	20.8	10.4	4.0
Max.	24.1	17.6	40.0	33.3	16.7	7.3
Range	10.9	7.9	36.0	12.5	6.3	3.3

5 重回帰分析による HRTF の生成

初期 HRTF の振幅レベルを目的変数, 13 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)).

$$y(s, \beta, f) = \sum_{i=1}^{13} a_i(\beta, f)x_i(s) + b(\beta, f) \quad (1)$$

ここで、 s, β, f, a_i, x_i, b は被験者、上昇角(0°-180°, 30°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。

各離散周波数において全目的変数を用いて算出した重相関係数の例(上昇角 0°)を Fig.4 に示す。20 kHz 以下の帯域では 0.6-0.7 程度の値をとる周波数が多いが、低い値となる周波数もみられた。

各上昇角の 20 kHz 以下の重相関係数の平均値(Table 3)は 0.57–0.63 であり、方向間に顕著な差はみられなかった。

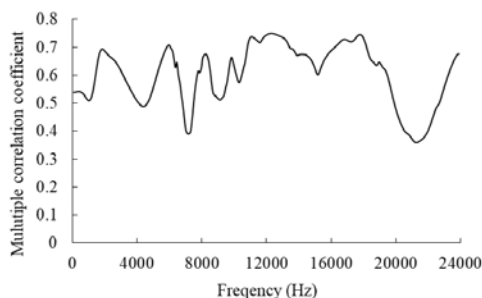


Fig.4 Multiple correlation coefficient as a function of frequency for vertical angle of 0°.

Table 3 Multiple correlation coefficient averaged over 93.75 Hz to 19,968.75 Hz.

0°	30°	60°	90°	120°	150°	180°
0.63	0.61	0.60	0.62	0.60	0.57	0.59

重回帰分析により推定した振幅スペクトルの例を Fig.5 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。ただし、推定値のノッチは実測値と比較してやや浅くなる傾向がみられる。

推定したスペクトルの精度を検証するため SD (spectral distortion,式(2))を算出した。

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

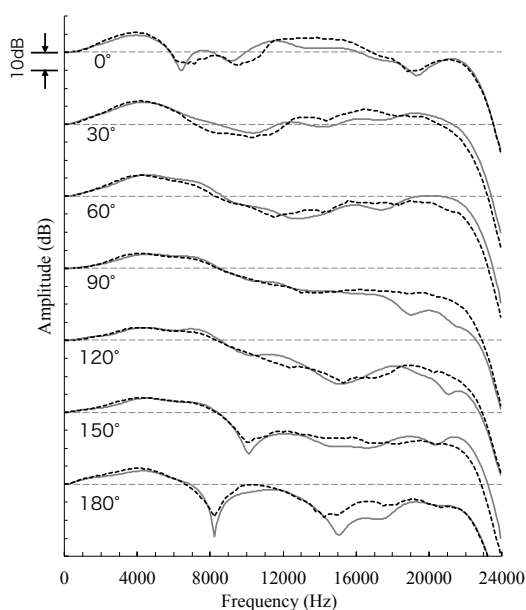


Fig.5 Examples of the measured early HRTFs (solid line) and the estimated HRTFs (broken line) in the upper

ここで $HRTF_m$, $HRTF_e$ はそれぞれ実測 HRTF と推定した HRTF, f は離散周波数を示す。

各上昇角における 54 耳の SD の最小値, 最大値, 平均値を Table 4 に示す。平均値は 3.1–3.8dB であり、スペクトルが平坦な 90°では低く、ノッチやピークが顕著な 0°や 180°では高くなる傾向がみられた。SD が大きくなる耳の特徴の分析とその対策が今後の課題である。

Table 4 SD between the measured early HRTF and the generated HRTF averaged over 54 ears, in decibel.

0°	30°	60°	90°	120°	150°	180°
3.8	3.4	3.6	3.1	3.7	3.6	3.7

6 ナイーブな被験者の HRTF 生成精度

交差検証法により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。異なる3名をナイーブな被験者とみなして9回の交差検証法を行った(3名×9回=27名)。

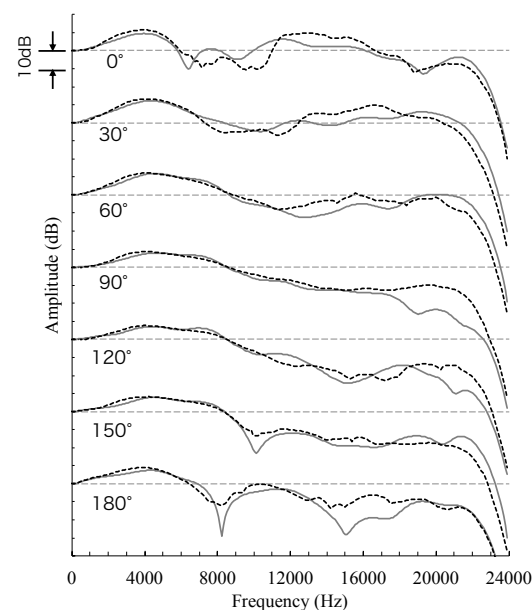


Fig.6 Examples of the measured early HRTFs (solid line) and the generated HRTFs (broken line) of a naïve subject.

生成したナイーブな被験者の HRTF の一例 (Fig.5 の重回帰分析と同じ耳)を Fig.6 に示す. ここではナイーブな被験者として耳介形状データのみから生成したため, Fig.5 の推定値と比較すると精度はやや低下した. しかし, 特徴的なノッチやピークは再現されている.

交差検証法で求めた全てのナイーブな被験者(27名(54耳))の SD の平均値を上昇角毎に Table 5 に示す. 平均値は 4.6–5.8dB であり, スペクトルが平坦な 90°では低く, ノッチやピークが顕著な 0°や 180°では高くなる傾向がみられた. SD が最大となった耳は, 重回帰分析で使用した耳よりも x_5 (最大耳長)が小さい傾向があった. 今後, さらに詳細な原因の分析と対策を進めたい.

Table 5 SD between the measured early HRTF and the generated HRTF averaged over 27 naive subjects (54 ears), in decibel.

0°	30°	60°	90°	120°	150°	180°
5.8	4.9	5.3	4.6	5.6	5.4	5.7

7 おわりに

正中面 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) 耳介形状から生成した HRTF には, 実測 HRTF にある特徴的なノッチやピークが観察された. ただし, 生成 HRTF のノッチは実測 HRTF と比較してやや浅くなる傾向がみられた.
- 2) 重回帰分析に用いた被験者における推定 HRTF と実測 HRTF の平均 SD は正中面 7 方向で 3.1–3.8dB であった.
- 3) ナイーブな被験者における生成 HRTF と実測 HRTF の平均 SD は正中面 7 方向で 4.6–5.8dB であった.

今後, さらに以下の取組みが必要である.

- ① 物理的な生成精度の向上. 特に生成誤差の大きい耳の原因分析と対策.
- ② 生成した HRTF を用いた音像定位実験による知覚面での精度の検証.
- ③ 耳介形状データの簡易取得法の確立.

参考文献

- [1] 飯田, 頭部伝達関数の基礎と 3 次元音響再生への応用, 63–70, コロナ社, 2017.
- [2] Morimoto and Ando, J. Acoust. Soc. Jpn. (E) 1, 167–174, 1980.
- [3] Bomhardt *et al.*, Proc. Meetings on Acoustics, 29, 050007, 2017.
- [4] Iida *et al.*, J. Acoust. Soc. Am., 136, 317–333, 2014.
- [5] Iida *et al.*, Appl. Acoust. 68, 835–850, 2007.
- [6] Iida and Ishii, Appl. Acoust., 129, 239–247, 2018.
- [7] 島崎と飯田, 音講論(春), 1585–1588, 2018.
- [8] Shaw, Acoustical features of the human external ear, Binaural and spatial hearing in real and virtual environments, Erlbaum, 25–47, 1997.
- [9] Takemoto *et al.*, J. Acoust. Soc. Am., 132, 3832–41, 2012.
- [10] Iida and Oota, Appl. Acoust., 139, 14–23, 2018.