

受聴者の耳介形状に基づいた頭部伝達関数の個人化精度の検証*

○石井要次, 西岡伸介, 飯田一博 (千葉工大)

1 はじめに

受聴者本人の頭部伝達関数(HRTF)を鼓膜上で再現することで任意の方向に音像を制御できることが広く知られている[1]. また, 上昇角知覚の手掛かりは, HRTF の 4kHz 付近に存在するピーク(P1), および 4kHz 以上に存在する最初のノッチ(N1)とその次のノッチ(N2)であることが知られている[2].

他人の HRTF を用いると, 前後誤判定や頭内定位などの現象がしばしば発生する[1]. しかしながら, すべての受聴者の HRTF を測定することは, 膨大な時間と無響室などの設備が必要であり, 現実的ではない.

一方, 耳介形状から受聴者の HRTF を推定することが考えられるが[3], HRTF がもつ全ての情報を推定することは困難である.

我々は, 測定や試聴なしに, 耳介形状から N1, N2, P1 周波数を推定し, 受聴者に適合する HRTF をデータベースから選出する方法について研究を進めている. 前報[4]では, 46 耳について, 耳介形状データ (Fig. 1) を説明変数, 正面方向の N1, N2, P1 周波数を目的変数とした重回帰分析を行った. その結果, 約 90% の確率で N1, N2, P1 周波数を弁別閾内の残差で推定できることを示した.

本報では, 重回帰分析に用いていない (ナイーブな) 被験者を用いて, 前報で提案した HRTF の個人化方法の精度を物理的・知覚的側面から検証した.

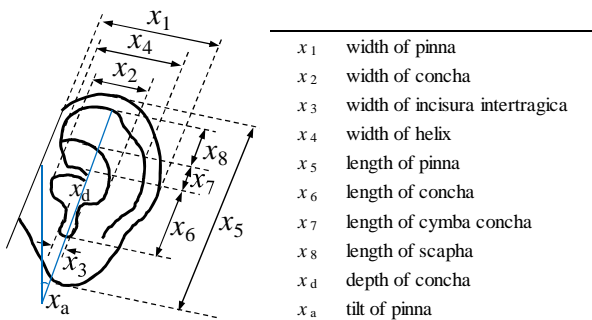


Fig. 1 Ten anthropometric parameters used in analysis

2 N1, N2, P1 周波数の推定精度の検証

重回帰分析に用いていない3名(TCY, YMM, OIS)の両耳について, 前報で分析した以下の2つの回帰モデル(A, B)による N1, N2, P1 周波数の推定精度を検証した.

モデル A : Fig. 1 に示す全ての耳介形状データ($x_1 - x_d, x_a$)を説明変数に用いる.

モデル B : 耳介写真から計測することができない耳甲介腔の深さ(x_d)以外を説明変数に用いる.

実測値と推定値の分布を Fig. 2 に示す. N1, N2, P1 周波数の推定値は, いずれのモデルにおいても, 概ね対角線上に分布しており, 良好な推定ができています.

Table 1 に残差を示す. 6 耳の残差の平均値は, N1 では, Model A と B それぞれ 418, 317 Hz, N2 では, 561, 518 Hz, P1 では 115, 134 Hz であった.

ここで, 正面方向における P1 周波数の弁別閾は, 約 0.3oct. [5]であることが示されている. また, N1, N2 周波数の弁別閾は, 約 0.1 - 0.2 oct. [6]であることから, 今回は 0.15oct. とした. その結果, 残差は, TCY 右耳の N1 を除き, すべて弁別閾内となった. なお, TCY 右耳の N1 の残差は, Model A, B それぞれ 0.16, 0.18oct. であった.

3 best-matching HRTF の選出

データベースから推定結果に最も近い HRTF (best-matching HRTF) を選出し, その有効性を音像定位実験により検証した.

3.1 選出方法

N1, N2, P1 周波数の推定値に最も近い HRTF をデータベースから選出するための物理評価指標として, NFD (Notch Frequency Distance) [6]を用いた. 推定値との NFD が最小となる HRTF をデータベースから探索する. P1 周波数の弁別閾は, 個人差[4]よりも大きいいため, P1 は HRTF 選出の尺度には用いない.

* Verification of accuracy of individualization of head-related transfer function based on anthropometry of the listener's pinna, by ISHII, Yohji, NISHIOKA, Shinsuke and IIDA, Kazuhiro (Chiba Institute of Technology).

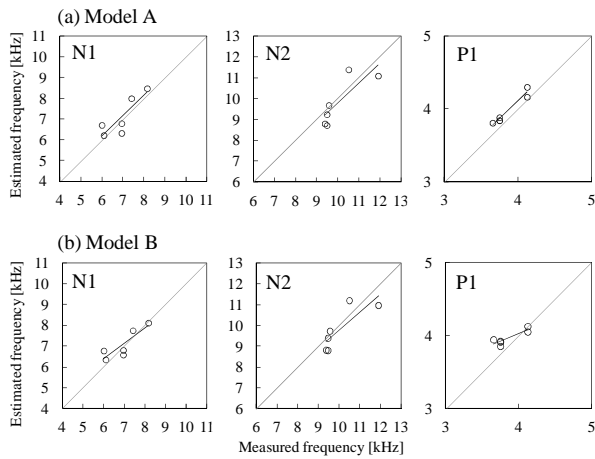


Fig. 2 Relation between the measured and estimated N1, N2, and P1 frequencies

Table 1 Residual error

model	absolute mean residual error		The number of pinnae of residual error is less than ind	
	[Hz]	[oct.]		
A: all params	N1	418	0.087	5 / 6
	N2	561	0.080	6 / 6
	P1	115	0.043	6 / 6
B: except x_d	N1	317	0.068	5 / 6
	N2	518	0.074	6 / 6
	P1	134	0.050	6 / 6

正面方向について本研究の HRTF データベースを用いて探索を行った。探索対象は正中面内 7 方向(0-180, 30° 間隔)のデータが備わる 112 耳である。今回は、左右の耳で個別に best-matching HRTF を選定した。

3.2 best-matching HRTF の物理的精度

Fig. 3 に被験者本人の HRTF と、回帰モデル A の推定結果に基づいて選出した best-matching HRTF の振幅スペクトルを示す。いずれの best-matching HRTF も、その N1,N2,P1 の構造的特徴が被験者本人の HRTF と類似していることが見てとれる。ただし、ピークやノッチのレベルは推定対象としていないため、この点については差異がある場合もある(例えば TCY 右耳の N2)。次に、被験者本人の HRTF と、best-matching HRTF の N1, N2, P1 周波数の差を求めた (Table 2)。ここでは、N1, N2, P1 周波数の推定誤差と、データベースから選出する際の誤差が重ね合わされて現れる。

まず、N1, N2 についてみると、平均誤差は、0.09, 0.08oct であった。最大誤差は、TCY 右耳の N1 で 0.15oct であった。P1 については、平均誤差は 0.06 oct., 最大誤差は 0.16oct. であった。N1, N2, P1 周波数の残差は、いずれの耳介においても弁別閾内となった。

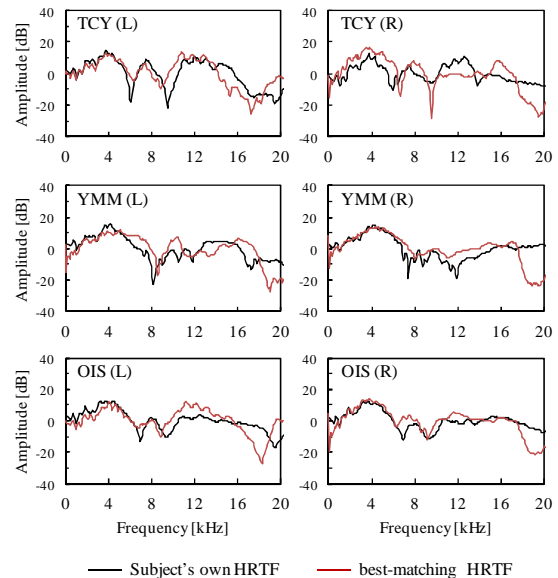


Fig. 3 Subject's own HRTF (black line) and selected best-matching HRTF (red line)

Table 2 Frequency differences of N1, N2 and P1 between the subject's own HRTFs and the best-matching HRTFs

Pinna		[Hz]			[oct.]		
		N1	N2	P1	N1	N2	P1
TCY	L	125	656	0	0.03	0.10	0.00
	R	656	0	188	0.15	0.00	0.07
YMM	L	375	1031	469	0.06	0.14	0.16
	R	656	656	0	0.12	0.08	0.00
OIS	L	188	563	281	0.04	0.09	0.10
	R	656	281	94	0.14	0.04	0.04
Average		443	531	172	0.09	0.08	0.06

3.3 best-matching HRTF の心理的精度

best-matching HRTF の有効性を検証するため、音像定位実験を行った。

3.3.1. 実験条件

被験者は TCY, YMM, OIS の 3 名である。実験は防音室で行った。実験システムは、ノート PC, オーディオインターフェース, アンプ, オープンエアヘッドホン, A/D コンバータ, イヤーマイクロホンで構成した。再生系伝達関数をキャンセルするため、実験前にヘッドホンからイヤーマイクまでの伝達関数を測定し、その逆特性を刺激に畳込んだ。音源信号は 200 - 17000 Hz のホワイトノイズで、以下の HRTF を畳込んだものを刺激とした。

- 1) 本人の実測 HRTF
 - 2) best-matching HRTF
 - 3) HATS(B&K, Type 4128C)の HRTF
 - 4) 他の被験者(1 人目)の HRTF
 - 5) 他の被験者(2 人目)の HRTF
- 目標方向は正中面 7 方向(0 - 180, 30°間隔)

である。best-matching HRTF は、今回は正面方向のみ得られるが、他の6方向では正面方向で選出したHRTFの提供者のものを用いた。なお、best-matching HRTF は、左右の耳で提供者が異なる場合があるため、ITDとILDが0となるよう微調整を施した。刺激のレベルは外耳道入口で63dB SPLで、継続時間は1.2秒(前後に0.1秒の立上がり/下がりを含む)とした。刺激はランダムな順に提示した。各刺激の提示回数は10回である。正中面を表す円と正面方向を表す矢印がノートPCの画面に表示され、被験者は、音像の上昇角をPCの画面上でクリックした。頭内定位した場合は、画面上にあるチェックボックスにチェックするように指示した。

3.3.2. 実験結果

A) 回答の分布

Fig. 4-6に各被験者の回答の分布を示す。まず、被験者TCY(Fig.4)では、本人のHRTF(a)では、120, 150°でやや前後誤判定がみられるが、概ね良好な定位精度を示している。best-matching HRTF(b)では、推定の対象とした0°では、本人のHRTFと同等の定位精度を示した。他の方向では、90°を除き、本人のHRTFと同等の定位精度を示した。HATSのHRTF(c)では、0, 180°では目標方向に定位したが、その他の方向では、60-120°には全く定位していない。YMMのHRTF(d)では、全体的にばらつきが大きい。OISのHRTF(e)では、0, 30, 180°で精度よく定位しているが、他の方向では、ばらつきが大きい。

被験者YMM(Fig.5)では、本人のHRTF(a)では0-90°で目標方向よりやや後方に定位する傾向がある。best-matching HRTF(b)では、推定の対象とした0°では、本人のHRTFとほぼ同等の定位精度を示している。他の方向でも、150°で上昇するものの、本人のHRTFと同様の定位精度を示している。HATS(c)のHRTFでは、0°には定位せず、すべての目標方向でばらつきが大きい。他の被験者のHRTF(d,e)では、180°では目標方向に定位したが、0°には定位せず、他の目標方向でも回答のばらつきが大きい。

被験者OIS(Fig.6)では、本人のHRTF(a)では、60°方向でやや後方に、120°ではやや前方に回答する傾向がある。best-matching HRTF(b)では、推定の対象とした0°では本人のHRTFと

同等の定位精度であった。その他の方向でも、150°でばらつくものの、概ね本人のHRTFと同等の定位精度である。HATSのHRTF(c)では、全体的に目標方向よりもやや後方に回答する傾向がある。TCYのHRTF(d)では、本人のHRTFとほぼ同等の定位精度である。YMMのHRTF(e)では、0, 90°でばらつきが大きく、全体的に上方に寄って回答する傾向がある。

B) 平均定位誤差

Fig. 7に被験者全員の回答についての平均定位誤差を示す。正面方向(a)では、被験者本人のHRTF(6°)とbest-matching HRTF(10°)では、良好な定位ができています。HATSおよび各被験者のHRTFの平均定位誤差は、40°以上と大きい。全方向の平均値(b)では、被験者本人のHRTFとbest-matching HRTFの平均定位誤差は23°で等しい。他の被験者のHRTFの平均定位誤差は、最小(OIS)で31°、最大(YMM)で58°であった。HATSでは、43°と大きな誤差となった。

C) 頭内定位

一部のHRTFで頭内定位が発生した。HATSのHRTFでは、被験者YMMが7%、TCYが27%の割合で、頭内に音像を知覚した。被験者YMMでは、OISのHRTFで18%、TCYのHRTFで9%の割合で、頭内に音像を知覚した。本人のHRTFとbest-matching HRTFでは頭内定位は発生しなかった。

4 結論

本研究では、耳介形状に基づいたHRTFの個人化方法の精度を検証した。その結果、以下のことが明らかになった。

- 1) 耳介寸法からN1,N2,P1周波数を推定し、データベースからbest-matching HRTFを選出した。best-matching HRTFと本人のHRTFのN1,N2,P1周波数の差異は、6耳中5耳で弁別閾内であった。
- 2) best-matching HRTFを用いて正中面内音像定位実験を行った結果、その定位精度は、被験者本人のHRTFと同等であった。また、頭内定位も発生しなかった。
- 3) ダミーヘッドや他人のHRTFを用いた場合は、顕著な定位誤差の低下や頭内定位が発生した。

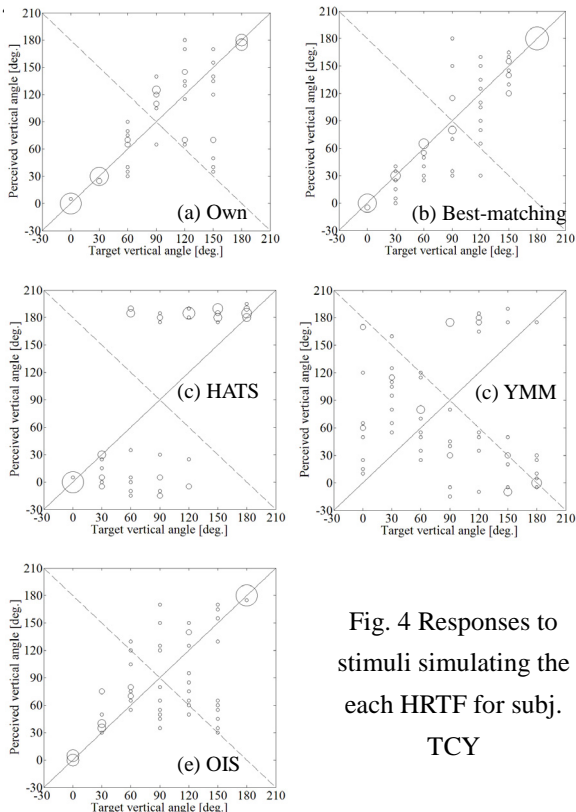


Fig. 4 Responses to stimuli simulating the each HRTF for subj. TCY

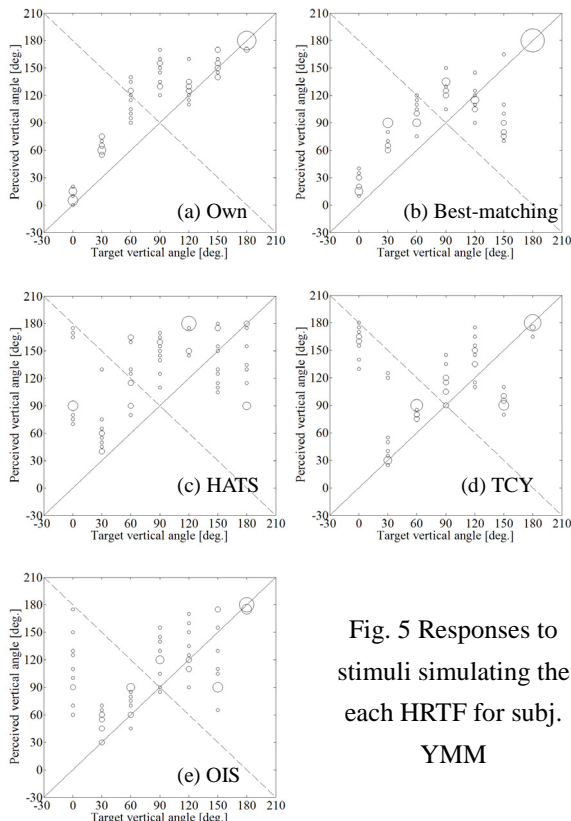


Fig. 5 Responses to stimuli simulating the each HRTF for subj. YMM

謝辞

本研究の一部は科研費(基盤研究(A) 22241040)により実施した。

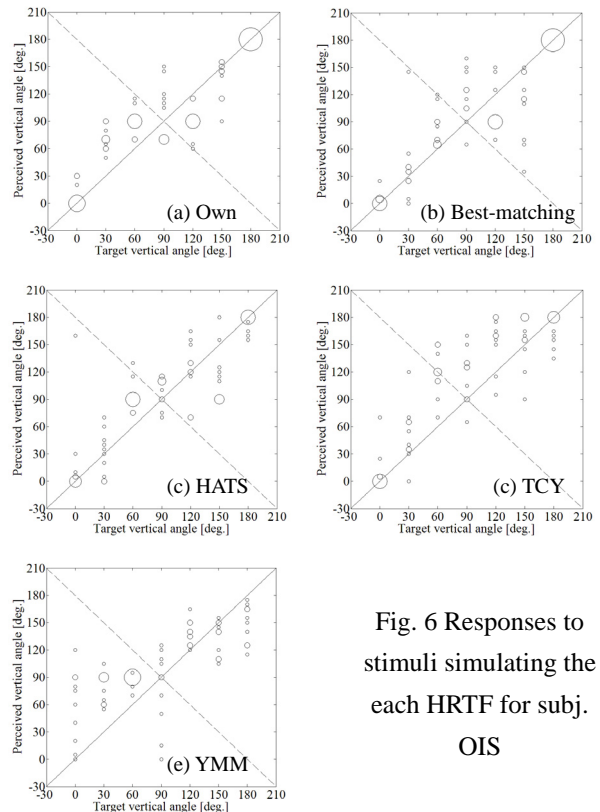


Fig. 6 Responses to stimuli simulating the each HRTF for subj. OIS

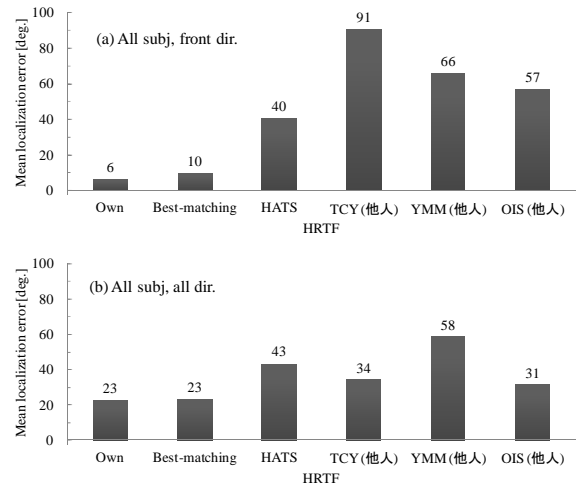


Fig. 7 Mean localization error.

参考文献

- [1] M. Morimoto and Y. Ando, J. Acoust. Soc. Jpn. (E), vol.1, pp. 167-174, 1980.
- [2] K. Iida *et al.*, Applied Acoustics, vol.68, pp. 835-850, 2007.
- [3] R. Sottek and K. Genuit, Proc. DAGA, 1999
- [4] 石井, 飯田, 音講論(秋), pp. 463-466, 2012
- [5] 西岡他, 音講論(秋), 2013
- [6] K. Iida and Y. Ishii, Proc. inter-noise 2011, 2011.