

正中面のPNPモデルによる全天空個人化頭部伝達関数の生成
- モデルの再構築と個人化プロセスの簡易化 -*

◎中村風香（千葉工大・院），飯田一博（千葉工大・先進工）

1. はじめに

3D 音響や音の VR を実現するために，頭部伝達関数(HRTF: Head-Related Transfer Function)の個人化方法の研究が活発に進められている[例えば 1-4]. われわれは聴取者自身で全天空の個人化 HRTF を生成できる PNP(Parametric Notch-Peak) HRTF モデルを提案している[5].

本研究では，生成した個人化 HRTF の定位精度の向上を目的として多数の HRTF データを用いて PNP モデルを再構築した. また，個人化プロセスをより簡単にしたアルゴリズムを考案し，それを搭載したアプリケーションを開発した. さらに，音像定位実験により提案法の有効性を検証した.

2. 個人化 HRTF 生成の全体像

本研究で提案する全天空の個人化 HRTF 生成方法は以下の(1)-(4)のプロセスで構成される(Fig. 1).

- (1) 正中面 7 方向の PNP モデルの構築
 - (2) 天頂方向の個人化 HRTF の生成
 - (3) 正中面の個人化 HRTF の生成
 - (4) 全天空の任意方向の個人化 HRTF の生成
- 各プロセスの詳細を 3-6 章で述べる.

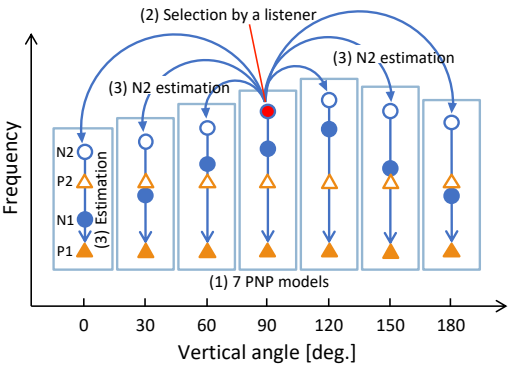


Fig. 1 Schematic diagram for generating the individualized HRTFs using seven PNP HRTF models.

3. 正中面 7 方向の PNP モデルの構築

正中面において実測 HRTF と同等の音像定位精度を提供する最小の HRTF の構成要素は，4 kHz 以上で最も周波数の低い 2 つのノッチ (N1,N2)と 2 つのピーク(P1,P2)であることが報告されている(Fig. 2)[6,7].

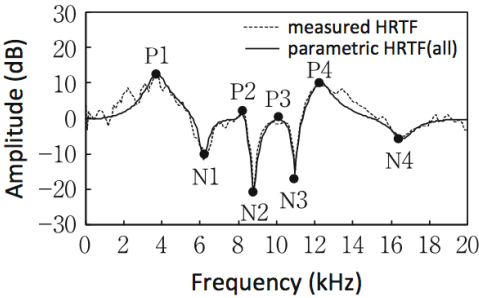


Fig. 2 An example of parametric HRTF.

この 4 つのノッチとピークは，それぞれ IIR のピーキングフィルタを用いることにより，中心周波数，レベル，尖鋭度(Q)の 3 つのパラメータで決定される. したがって， $4 \times 3 = 12$ パラメータを適切に設定すれば各聴取者に適した HRTF が生成できる.

PNP モデルは，N2 周波数を唯一の独立変数として他のパラメータを従属的に表現する[5]. パラメータの設定方法を 3.1 節および 3.2 節に示す.

3.1 N1・P1・P2 周波数の設定

正中面 7 方向において，成人 118 耳介を用いて各方向の N2 周波数を独立変数，N1,P1,P2 周波数を従属変数とする回帰分析を行った. 結果の一例として Fig. 3 に正面方向の N2 周波数と N1 周波数の関係を示す.

回帰式により推定した N1, P1, P2 周波数の平均残差を Table 1 に示す. いずれの方向においても平均残差は弁別閾 (0.1-0.2 oct.)[8]よりも小さかった. また，残差が弁別閾内に収まる耳介は，N1 周波数では 79-98%，P1 は 100%，P2 は 96-100%であった.

* Generation of the individualized HRTFs in the upper hemisphere using parametric notch-peak models in the median plane - Reformulation of the model and simplification of the individualization process -, by NAKAMURA, Fuka and IIDA, Kazuhiro (Chiba Institute of Technology).

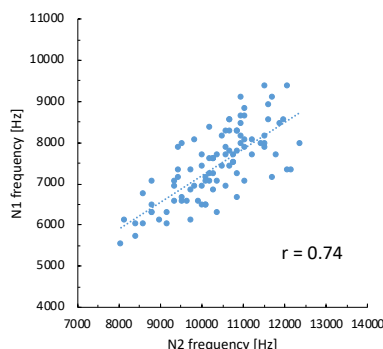


Fig. 3 Relationship between the N2 frequency and the N1 frequency for the front direction.

Table 1 Mean residual errors for N1, P1, and P2 estimation in the upper median plane [oct.].

Vertical angle (deg.)	N1	P1	P2
0	0.09	0.06	0.09
30	0.09	0.06	0.11
60	0.09	0.07	0.10
90	0.06	0.07	0.10
120	0.05	0.06	0.09
150	0.07	0.07	0.09
180	0.08	0.08	0.09

3.2 レベルおよびQの設定

各方向のノッチレベルは118耳介のHRTFの第一四分位数，ピークレベルは第三四分位数を用いた．ノッチとピークのレベルを平均値ではなく，第一四分位数と第三四分位数とした理由は，聴取者がノッチとピークを検知しやすくするためである．Qは平均値を用いた．

4. 天頂方向の個人化 HRTF の生成

PNP モデルを用いて全天空の個人化 HRTF を生成するうえで，少なくとも1方向については聴取者自身が試聴を通してN2周波数の調整・選定を行う必要がある．本研究ではそれを天頂方向とした．その理由は，天頂はN2周波数の違いによる音像方向の弁別が容易で，試聴により各聴取者に適したN2周波数を判断しやすい方向であるためである．

天頂方向の118耳介の実測N2周波数の分布範囲を求めた．最小値と最大値は，それぞれ11813Hzと16219Hzであった．この分布範囲を1/12oct.(0083 oct.)間隔で9種類のカテゴリに分けた．先述のようにN2周波数の弁別閾は0.1-0.2 oct.であるため，カテゴリの間隔は弁別閾以下である．天頂方向のPNPモデルを用いて，各カテゴリのN2周波数に対応した9種類のHRTFを生成した．

聴取者は水平面と横断面の音像移動により，

9種類のHRTFを聴き比べて最も本人に適合するHRTFを選択する．適合するカテゴリでは，水平面移動では円軌道を描き，横断面移動では天頂を通る．一方，適合していないカテゴリではいずれも軌道は頭内を通過する．

5. 正中面の個人化 HRTF の生成

5.1 正中面6方向の個人化 HRTF の生成

4章で生成した天頂方向のHRTFから上半球正中面の他の6方向(上昇角0°, 30°, 60°, 120°, 150°, 180°)の個人化HRTFを生成する．

5.1.1 N2周波数の推定

天頂方向のN2周波数から他の6方向のN2周波数を推定する．

118耳介を用いて天頂方向のN2周波数を独立変数，正中面の他の6方向のN2周波数を従属変数とする回帰分析を行った．

回帰式により推定した6方向のN2周波数の平均残差，および残差が弁別閾に含まれる割合をTable2に示す．平均残差は0.03-0.07 oct.であり，いずれの方向においても弁別閾よりも小さかった．また，残差が弁別閾内に収まる耳介は全体の87-100%であった．

Table 2 Mean residual errors for estimated N2 frequency and ratio of pinna of which residual is within jnd.

Vertical angle (deg.)	Mean residual error (oct.)	Ratio of pinna of which residual is within jnd
0	0.07	0.87
30	0.06	0.90
60	0.04	0.99
120	0.03	1.00
150	0.05	0.99
180	0.06	0.95

5.1.2 N1・P1・P2周波数の推定

正中面6方向のN1,P1,P2周波数は3章で構成した各方向のPNPモデルを用いてN2周波数から推定した．

5.1.3 レベルおよびQの設定

3.2節で述べたように，各方向のノッチレベルは成人118耳のHRTFの第一四分位数，ピークレベルは第三四分位数を用いた．Qは平均値を用いた．

5.2 正中面の任意方向の個人化 HRTF の生成

正中面7方向の個人化HRTFのパラメータの線形補間により正中面の任意の方向のパラメータを推定し，個人化HRTFを生成した．

5.3 後方知覚の手掛かり(P0)の付加

最近の研究により[9,10], 1kHz 付近で後方の HRTF は正面よりも平均で 5.6 dB 卓越し, この卓越帯域に対応するピーク(以降 P0 と呼ぶ)を追加した HRTF モデルは従来モデルよりも後方の音像定位精度が有意に向上することを示した.

P0 の周波数および卓越量の個人差は小さいことから, 正中面の上昇角 120-180°の PNP モデルに固定値の P0 を加えた.

6. 全天空の任意方向の個人化 HRTF の生成

上半球正中面の個人化 HRTF に側方角に応じた両耳間差情報を加えて[11], 全天空の任意の方向の個人化 HRTF を生成した(Fig. 4).

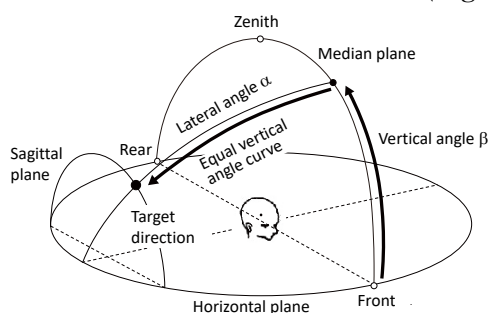


Fig. 4 Sound image localization for arbitrary directions in the upper hemisphere using the individualized HRTFs in the upper median plane and interaural differences.

7. 個人化 HRTF の音像定位精度の検証

2 章から 6 章までに述べたアルゴリズムを搭載した全天空個人化 HRTF 生成アプリを MATLAB で開発した(Fig. 5). このアプリを用いて聴取者自身が個人化 HRTF を生成し, 音像定位実験により定位精度を検証した.

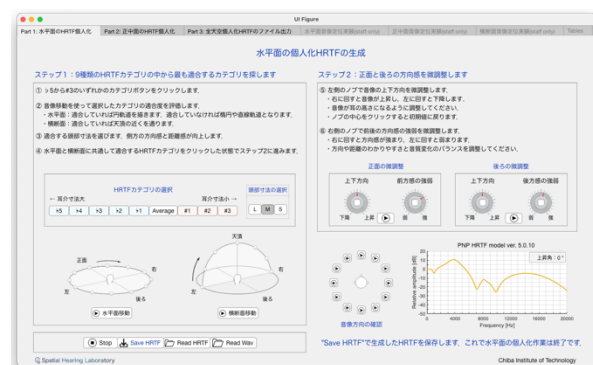


Fig. 5 GUI of the application software of PNP model for HRTF individualization.

7.1 水平面の音像定位実験

7.1.1 実験方法

目標方向は右半分の水平面内(方位角: 0-180°, 30°間隔)の 7 方向である. 音源信号は

広帯域白色雑音(上限周波数 20kHz)で, 提示時間は 1s である. 音源信号に個人化 HRIR を畳み込んだ刺激をヘッドホン(DT990 PRO, beyerdynamic)により被験者に提示した. ヘッドホン伝達関数は補正していない. 各刺激はランダムな順に 10 回ずつ提示した. 被験者は各刺激を何回でも聴くことができるが, 回答後はその刺激を聴くことはできない. 被験者は成人 3 名(男性 2 名, 女性 1 名)で, マッピング法により音像の方位角と仰角を回答した.

7.1.2 実験結果

各被験者の方位角の回答を Fig. 6 に示す. いずれの被験者においても目標方向付近に回答したが, 被験者 2 は目標方位角 30-120°において回答にややばらつきがみられ, 被験者 3 では目標方位角 120°において回答が後方に寄った. 仰角についてはいずれの被験者においても目標方向付近に回答した.

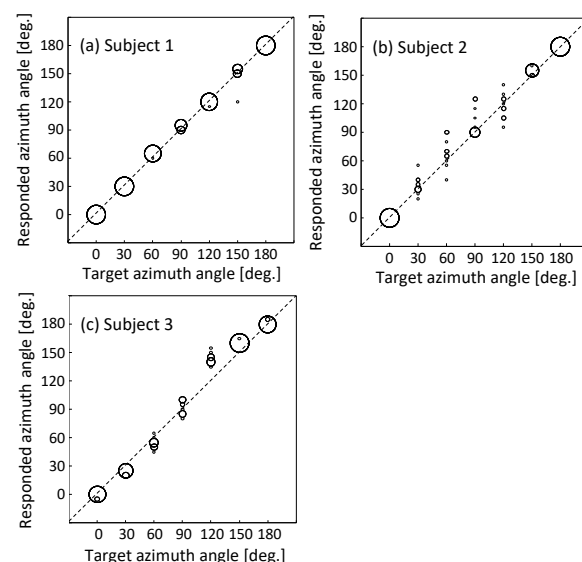


Fig. 6 Responded azimuth angle to the subject's individualized HRTFs in the horizontal plane.

7.2 正中面の音像定位実験

7.2.1 実験方法

実験方法は水平面内の場合と同様である. ただし, 目標方向は上半球正中面内(上昇角: 0-180°, 30°間隔)の 7 方向であり, 被験者は音像の上昇角を回答した.

7.2.2 実験結果

各被験者の回答を Fig. 7 に示す. いずれの被験者においても目標方向付近に回答したが, 被験者 3 では目標上昇角 30°, 90°で 60°に寄り, 目標上昇角 120°, 150°では後方に寄る傾向がみられた.

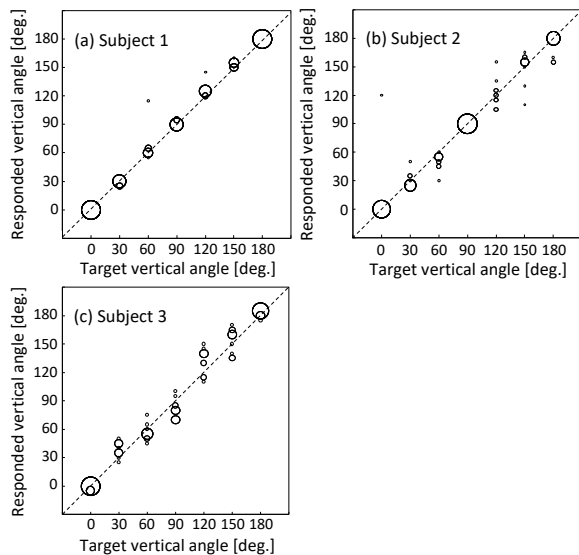


Fig. 7 Responded vertical angle to the subject's individualized HRTFs in the median plane.

7.3 横断面の音像定位実験

7.3.1 実験方法

実験方法は水平面内の場合と同様である。ただし、目標方向は上半球横断面内(側方角: -90° から $+90^{\circ}$, 30° 間隔)の 7 方向であり、被験者は音像の側方角と上昇角を回答した。

7.3.2 実験結果

被験者が回答した側方角、上昇角を Fig. 8, 9 に示す。側方角については、いずれの被験者も目標方向付近に回答したが、被験者 2 では目標方向 30° 付近で、被験者 3 では全体的に正中面側に寄る傾向がみられた。

上昇角については、被験者 1, 2 は目標方向 (90°) 付近に回答したが、被験者 3 は -60° , -30° ではやや前寄りに、 30° , 60° ではやや後ろ寄りに回答した。

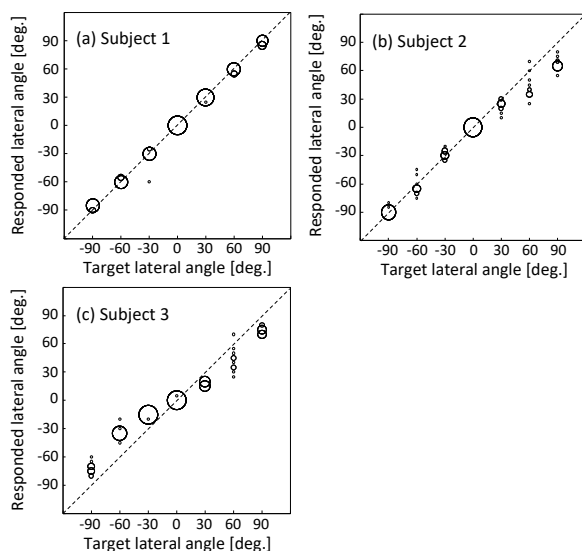


Fig. 8 Responded lateral angle to the subject's individualized HRTFs in the transverse plane.

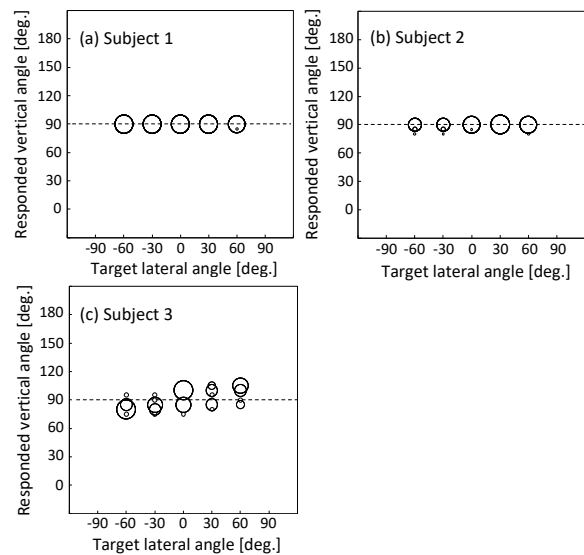


Fig. 9 Responded vertical angle to the subject's individualized HRTFs in the transverse plane.

8. おわりに

本研究では、多数の HRTF を用いて正中面 7 方向の PNP HRTF モデルを再構築し、全天空の任意の 3 次元方向の個人化 HRTF を生成するアプリケーションを開発した。音像定位実験の結果、水平面、正中面、横断面のいずれにおいても良好な音像定位が可能であることが示された。

現在、多数の被験者を用いて音像定位精度と操作性の検証を進めている。

参考文献

- [1] Kistler and Wightman, J. Acoust. Soc. Am. 91,1637-47, 1992.
- [2] Middlebrooks and Green, J. Acoust. Soc. Am. 92, 597-599, 1992.
- [3] Bomhardt *et al.*, Proc. of meetings on acoustics, 2016.
- [4] Miccini and Spagnol, IEEE 5th VR Workshop on Sonic Interactions in Virtual Environments, Mar. 2020.
- [5] 中村と飯田, 電子情報通信学会技術研究報告, EA2022-82, 2023.
- [6] Iida *et al.*, Appl. Acoust. 68, 835-850, 2007.
- [7] Iida *et al.*, Appl. Acoust. 129, 239-247, 2018.
- [8] 飯田, 頭部伝達関数の基礎と 3 次元音響再生への応用, 66-67, コロナ社, 東京, 2017.
- [9] 中村と飯田, 日本音響学会聴覚研究会資料, (2023 年 8 月)
- [10] 中村と飯田, 音講論(秋), 2-11-1, 2023
- [11] Morimoto *et al.*, Acoust. Sci. & Tech. 24, 267-275, 2003.