

Parametric Notch-Peakモデルによる頭部伝達関数の個人化方策の検証*

◎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

3次元音響や音の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[1-4]. われわれは、ヒトの方向知覚の手掛かりに着目した Parametric Notch-Peak HRTF モデル(PNP モデル)を用いて全天空の個人化 HRTF を生成するアルゴリズムを提案し(Fig.1), アプリケーションソフトを開発している[5]. このアプリでは聴取者自身が天頂方向のN2 周波数を個人最適化すると, 全天空の個人化 HRTF が生成される.

天頂方向のN2 周波数の個人最適化についてはさまざまな方法が考えられるが, 3次元音響を実用化するには, 専門知識をもたない一般の聴取者が操作可能な方法にする必要がある. われわれは, いくつかの試作を経て“9種類のN2 周波数カテゴリから聴取者自身がフィットするカテゴリを選択する”方法が操作の簡便性と定位精度のバランスがよいと考えている.

本稿では, PNP モデルによる HRTF 個人化手順の簡便性を評定尺度法により評価する. また, 音像定位実験により, 生成した個人化 HRTF の音像定位精度を検証する.

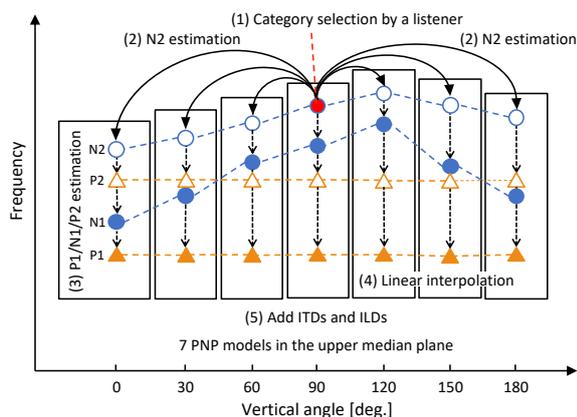


Fig. 1 Schematic diagram for generating individualized HRTFs in the upper hemisphere using seven PNP HRTF models in the upper median plane.

2. 個人化 HRTF 生成手順の評価

MATLAB®の App Designer で開発した個人化 HRTF 生成アプリを用いて, 個人化手順を評価した. 比較のため, N2 周波数が9 カテゴリのアプリの他に, 連続しない4 カテゴリを除外して5 カテゴリ構成としたアプリも評価した. ハードウェアはPC(MacBook Air, Apple), ヘッドホンアンプ(HA-P50SE, TEAC), ヘッドホン(DT990 PRO, beyerdynamic)を用いた.

2.1 個人化 HRTF の生成手順

以下の手順で個人化 HRTF を生成した.

(1) 聴取者自身が9種類もしくは5種類のN2 周波数カテゴリから適合するものを選定した. アプリ画面において, あるN2 周波数カテゴリ(Fig.2(a))と音像移動ボタン(Fig.2(b))をクリックすると, 水平面, 横断面, 正中面のHRTFが生成され, 白色雑音と畳み込まれて各面での移動音像がヘッドホンから再生される. 聴取者は音像の軌跡を以下の点に着目して評価し, 最も適合するカテゴリを選定した.

- ・横断面と正中面移動: 適合すれば音像は天頂を通る. 適合しなければ頭内を通る.

- ・水平面移動: 適合すれば音像が円周上を動く. 適合しなければ, 音像は左右に長い楕円や, 左右耳を結び頭内を通る直線上を動く.

(2) 次に, 3種類の頭部寸法カテゴリ(S, M, L)から適合するものを選定した(Fig.2(c)). 適合するカテゴリでは, 音像が目標方向に近づくのみならず, 音像距離が増大し頭外感が向上する.

(3) 選定したN2 周波数カテゴリと頭部寸法カテゴリを用いて, 水平面8方向(方位角: 0-180°, 45°間隔)と上半球正中面3方向(上昇角: 45, 90, 135°)の計11方向の個人化 HRTF を生成した. 音源の方位角と仰角を数値またはスライダで設定すると(Fig.2(d)), その方向のHRTFが生成され, 音源と畳み込んだ音が再生される. 本実験では以下の2種類の方法で生成した.

- ・生成方法 A: Fig.2(d)に目標の方位角と仰角を入力して個人化 HRTF を生成した.

- ・生成方法 B: Fig.2(d)に目標の方位角と仰角

* Verification of a strategy for generating individualized head-related transfer functions using Parametric Notch-Peak HRTF model, by NAKAMURA, Fuka and IIDA, Kazuhiro (Chiba Institute of Technology).

を入力して試聴し、音像方向が目標方向と異なる場合は、目標方向と一致するように方位角と仰角を修正して個人化 HRTF を生成した。

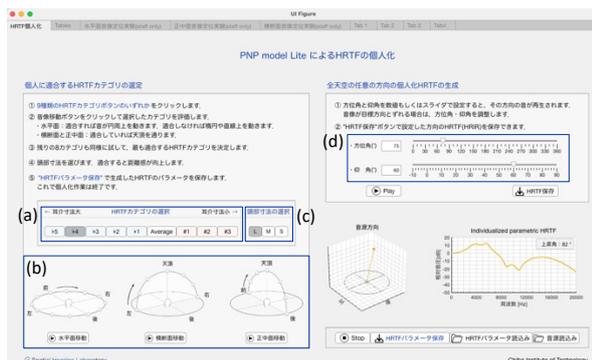


Fig. 2 Application software for HRTF individualization using Parametric Notch-Peak HRTF model.

2.2 評価方法

被験者は以下の項目について評価した。

Q1. 適合する N2 周波数カテゴリはあったか

Q2. 適合する頭部寸法カテゴリはあったか

Q3. 水平面の目標方向に知覚する HRTF を生成できたか

Q4. 正中面の目標方向に知覚する HRTF を生成できたか

Q5. アプリは使い易かったか

Q6. HRTF の個人化は容易に行えたか

Q7. 9 カテゴリと 5 カテゴリのどちらがより HRTF の個人化が容易に行えたか

被験者は 5 段階の評定尺度法により、-2.0 から 2.0 まで 0.1 間隔の目盛が記された数直線にペンで印をつけて回答した(Fig.3)。ここで、-2.0 は“不同意”，-1.0 は“やや不同意”，0.0 は“どちらともいえない”，1.0 は“やや同意”，2.0 は“同意”とした。被験者は正常な聴覚をもつ成人男女 18 名(男性 11 名，女性 7 名)である。

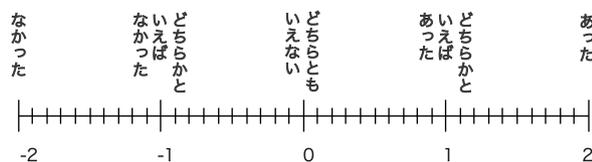


Fig. 3 Example of response sheet (for Q1).

2.3 評価結果

Q1 と Q2 の評価結果を Fig.4 に示す。Q1 の平均値は 9 カテゴリでは 1.3，5 カテゴリでは 1.4 であった。双方とも -1.0 と評価した被験者が 1 名いた。Q2 の平均値は 9 カテゴリ，5 カテゴリとも 1.3 であった。双方とも 0.0 と評価

した被験者が 1 名いた。

Q3 の評価結果を Fig.5 に示す。カテゴリ数に関わらず 90, 270°では評価が高く，平均値は 9 カテゴリでは双方ともに 1.8 であった。5 カテゴリでは 1.8 と 1.5 であった。一方，0°では 9 カテゴリで 0.9，5 カテゴリでは 0.7 であり，評価は低かった。

Q4 の評価結果を Fig.6 に示す。9 カテゴリと比較すると 5 カテゴリは 0, 45, 90°で評価が低く，平均値は 0.7, 0.6, 0.8 であった。

Q5 と Q6 の評価結果を Fig.7 に示す。Q5 の平均値は 9 カテゴリ，5 カテゴリともに 1.8 であり高い評価が得られた。最小値も双方ともに 1.0 であった。Q6 の平均値は 9 カテゴリでは 1.2，5 カテゴリでは 1.3 であり，HRTF の個人化は”やや容易に行えた”と”容易に行えた”の間の評価であった。ただし，9 カテゴリでは -1.0，5 カテゴリでは -0.5 と評価した被験者がそれぞれ 1 名いた。

Q7 の評価結果を Fig.8 に示す。ここでは -2.0 を“5 カテゴリ”，2.0 を“9 カテゴリ”として回答した。平均値は -0.4 であり，“やや 5 カテゴリ”と“どちらともいえない”の間の評価であった。ただし，-2.0 と 2.0 のそれぞれに回答した被験者がおり，評価が分かれた。

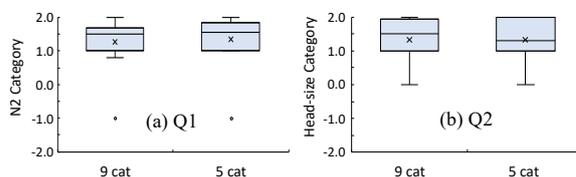


Fig. 4 Box plots for (a) Q1 and (b) Q2. x denotes mean value.

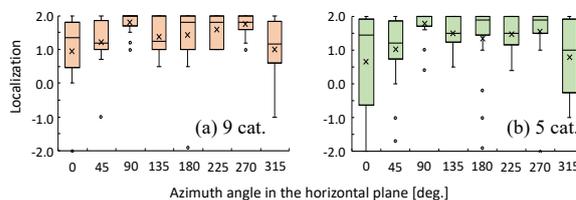


Fig. 5 Box plots for Q3. (a): nine categories and (b): five categories.

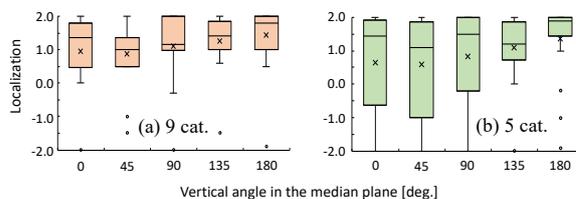


Fig. 6 Box plots for Q4. (a): nine categories and (b): five categories.

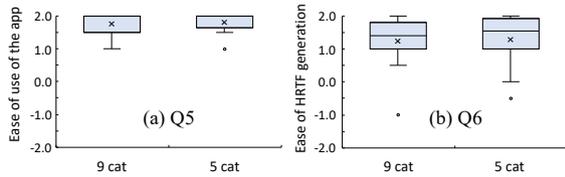


Fig. 7 Box plots for (a) Q5 and (b) Q6.

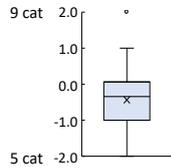


Fig. 8 Box plots for Q7.

3. 個人化 HRTF の音像定位精度の検証

アプリで生成した個人化 HRTF の音像定位精度を検証した。音像定位実験は HRTF 作成から少なくとも 36 日以上経過してから実施した。使用したハードウェアは 2 章と同様である。

3.1 水平面内の音像定位(9 カテゴリ)

3.1.1 実験方法

9 種類の N2 周波数カテゴリから生成した以下の個人化 HRTF を用いた。目標方向は水平面内の 8 方向(方位角: 0-315°, 45°間隔)である。

- ・個人化 HRTF(A): 生成方法 A による HRTF
- ・個人化 HRTF(B): 生成方法 B による HRTF
- ・本人実測 HRTF(0, 90, 180, 270°の 4 種類)

音源信号は広帯域白色雑音(20Hz-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。各刺激をランダムな順に 10 回ずつ提示した。被験者は成人 8 名(男性 3 名, 女性 5 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。実験は防音室で実施した。

3.1.2 実験結果

全被験者の個人化 HRTF(A),(B)と本人 HRTF に対する方位角の回答を Fig.9 に示す。比較のため、本実験に先立って無響室で実施した実音源(0-180°, 30°間隔)に対する回答を併せて示す。

個人化 HRTF(A),(B)はともに 0, 90, 180°では本人 HRTF および実音源と概ね同様の回答分布であった。270°では個人化 HRTF(A)は本人 HRTF と概ね同様の回答分布であったが、個人化 HRTF(B)では(A)よりも回答がばらついた。45, 135, 225, 315°ではやや回答がばらついた。個人化 HRTF(A)と(B)の間には顕著な差はみられなかった。

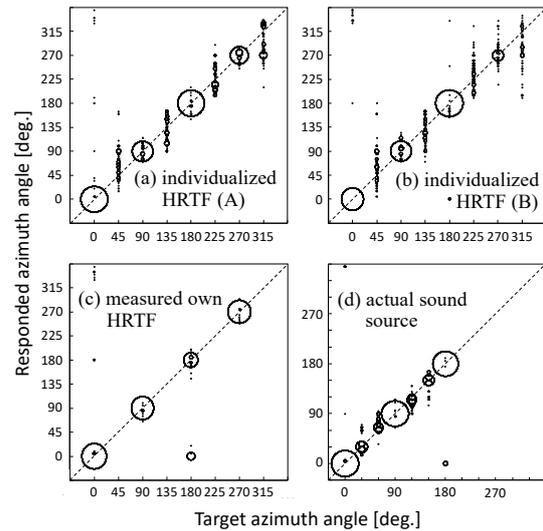


Fig. 9 Responded azimuth angle. (a) and (b): individualized HRTFs, (c): measured own HRTFs, and (d): actual sound sources in the horizontal plane.

3.2 正中面内の音像定位(9 カテゴリ)

3.2.1 実験方法

9 種類の N2 周波数カテゴリから生成した個人化 HRTF を用いて音像定位実験を実施した。目標方向は上半球正中面内の 5 方向(上昇角: 0-180°, 45°間隔)である。被験者は音像の上昇角を回答した。その他の実験方法は 3.1.1 項と同様である。

3.2.2 実験結果

全被験者の個人化 HRTF(A),(B)と本人 HRTF に対する上昇角の回答を Fig.10 に示す。比較のため、本実験に先立って無響室で実施した実音源(0-180°, 30°間隔)に対する回答を併せて示す。

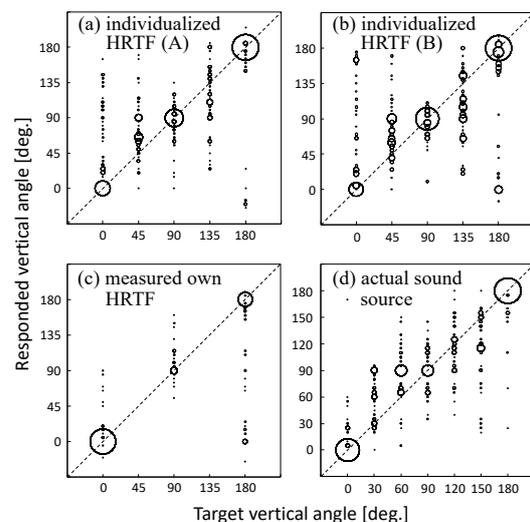


Fig. 10 Responded vertical angle. (a) and (b): individualized HRTFs, (c): measured own HRTFs, and (d): actual sound sources in the median plane.

個人化 HRTF(A),(B)はともに 90, 180°では本人 HRTF および実音源と概ね同様の回答分布であった。しかし, 0°では音像が上昇する場合があった。45, 135°では実音源よりもやや大きなばらつきが観察された。個人化 HRTF(A)と(B)の間には顕著な差はみられなかった。

3.3 9カテゴリと5カテゴリの比較

3.3.1 実験方法

8名中4名の被験者(男性1名,女性3名)は9カテゴリで選択したN2周波数カテゴリが5カテゴリにはなかった。5カテゴリの中からカテゴリを選択し直して生成した個人化 HRTF を用いて水平面および正中面で音像定位実験を実施した。その他の実験方法は3.1.1項と同様である。

3.3.2 実験結果

5カテゴリから生成した個人化 HRTF に対する回答を Fig.11(a),(b)に示す。生成方法はAである(2.1節)。比較のため, この4名の被験者の9カテゴリから生成した個人化 HRTF に対する回答を(c),(d)に示す。

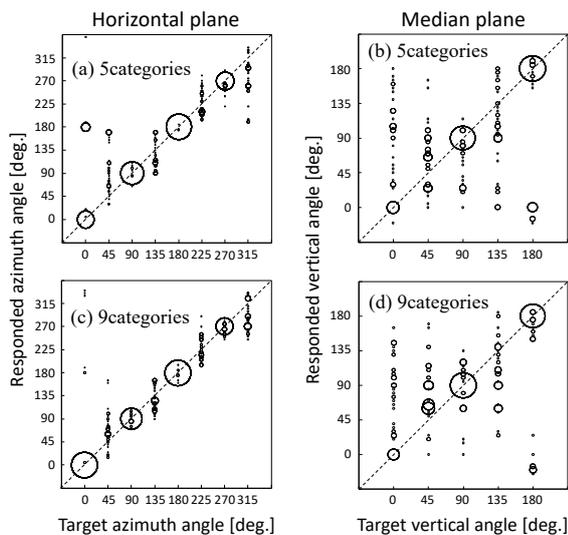


Fig. 11 Responses to the individualized HRTFs. (a) and (b): individualized HRTFs generated from five N2 frequency categories, (c) and (d): individualized HRTFs generated from nine N2 frequency categories.

水平面の回答分布(a),(c)では, 5カテゴリは9カテゴリとほぼ同様な傾向であったが, 0°では前後誤判定が増加し, 45, 135, 315°ではばらつきが増加した。正中面の回答分布(b),(d)では, 5カテゴリは9カテゴリと比較してばらつきや前後誤判定が増加する傾向がみられた。

水平面, 正中面ともに9カテゴリの方が音像

定位精度は高い傾向がみられた。ただし, Fig.8に示したように, 5カテゴリの中に適合するカテゴリが存在する場合は操作性のよい5カテゴリが好ましい。どちらが妥当なのかは応用先によると言える[6]。

4. おわりに

本稿では, PNP モデルを用いた HRTF 個人化手順の簡便性の評価実験と音像定位実験を実施し, 以下のことを示した。

(1) 目標方向に知覚する HRTF の生成については, 水平面, 正中面ともに“ややできた”と“できた”間の評価であった。しかし, 正中面の前半分においては, “どちらでもない”と“ややできた”の間であった。

(2) アプリは“使い易かった”に近い評価であった。

(3) HRTF の個人化は, “やや容易に行えた”と“容易に行えた”の間であった。

(4) 生成した水平面および正中面の個人化 HRTF の音像定位精度は, 本人 HRTF および実音源とほぼ同等であった。ただし, 正面方向においては音像が上昇する場合があった。

(5) 生成方法 A と B には顕著な差はみられず, N2 周波数カテゴリと頭部寸法カテゴリの選択だけでよいと言える。

(6) 9カテゴリと5カテゴリを比較すると音像定位精度では9カテゴリが, 個人化の容易さでは5カテゴリが上回った。

現在, 正面方向の音像定位精度を実音源と同程度に向上する方法について検証を進めている。

謝辞

学内外の多数の方にアプリの評価者および音像定位実験の被験者としてご協力いただきました。記して感謝いたします。

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, 音講論(秋), 269-272, 2023.
- [6] 飯田, 音講論(春), 3-6-3, 2024.