

頭部伝達関数の個人化に向けて*

- 聴覚の方向知覚の手掛りに基づいたアプローチ -

○飯田一博（千葉工大・工） 森本政之（神戸大院・工学研）

1 はじめに

ヒトは音の方向を知覚する際に頭部伝達関数(HRTF)に含まれる情報を利用しており、本人の HRTF を再現すると正確な音像定位ができる[1]. しかし、HRTF には頭部や耳介形状に起因する顕著な個人差が存在するため、他人の HRTF を用いると、音像の前後・上下方向の誤判定や、3次元音場再生における臨場感の低下などの問題が発生する. HRTF を利用した音像制御や音場再生が、長い研究の歴史をもつにもかかわらず、真の意味で実用化に至らない最大の理由は、HRTF の個人差を克服できていないからであろう.

HRTF の個人差を解決する方法の1つとして、“標準 HRTF”の作成が検討され、種々のダミーヘッドが開発されてきた[2]. しかしながら、耳介形状の平均値を用いて作成されたダミーヘッドは、結果的には、誰とも一致しない HRTF の持ち主となってしまった.

この問題の別の解決方法として、多くの受聴者に適用できるロバストな HRTF をもつ被験者を(何らかの方法で)選んで、その HRTF を利用するという考え方があつた. しかし、この方法で適用可能な受聴者の割合は決して大きいとは言えない. また、どの受聴者に適用可能なかを予想できないという問題も抱えている.

個人差を解決する最も本質的な方法は、受聴者それぞれに適合する HRTF を提供すること、つまり HRTF の個人化であろう. しかし、これを実現するには大きな課題が2つある. 第1の課題は、HRTF は振舞いが複雑で、しかも個人差が大きいことから(Fig.1)、そのままでは個人差を簡潔に記述することができないということである. HRTF に含まれる本質的な情報を抽出し、その情報の個人差で議論する必要がある.

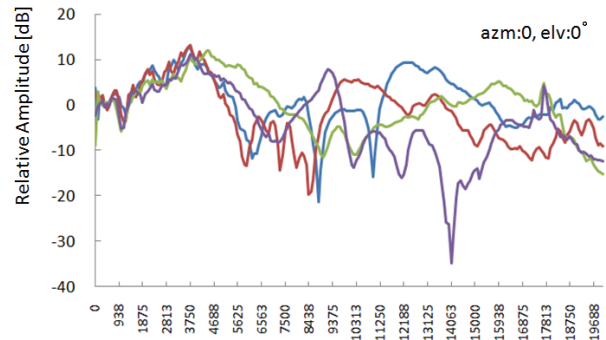


Fig.1 HRTFs of four subjects for vertical angle of 0 degree in the median plane.

また、予めデータベースに蓄えられた HRTF から本質的な情報を抽出して、それらの個人差を記述できたとしても、受聴者ひとりひとりの HRTF を測定したり数値計算したりすることはできないから、受聴者の HRTF (もしくはその本質的な情報)を推定する方法が必要である. これが第2の課題である.

これらの課題を解決し、HRTF の個人化(個人適応)を実現するために、著者は以下の3つのテーマで研究を進めている.

1) HRTF に含まれる方向知覚の cue の解明

複雑な振舞いをする HRTF の個人差を記述するために、HRTF の本質、つまり方向知覚の手掛り (cue) を解明し、cue の個人差を記述する.

2) 個人の cue の推定方法

一般に受聴者個人の HRTF を測定することは困難であるため、個人の方向知覚の cue を簡便に推定する方法を構築する.

3) HRTF の個人差の物理評価指標

ある HRTF がその受聴者に適合するか否かを cue に基づいて判定できる指標を提案する.

本稿では、これら3点について、これまでの研究成果と今後の展望を紹介する.

* Toward the individualization of the Head-Related Transfer Functions – An approach based on the cues for the sound localization of a human being - by IIDA, Kazuhiro (Chiba Institute of Technology), and MORIMOTO, Masayuki (Kobe Univ.).

2 ヒトの方向知覚の cue

ヒトの方向知覚の cue は以下の2つであることが知られている。

- 1) 両耳間差情報(ILD,ITD) : 左右方向
 - 2) スペクトル情報 : 前後・上下方向
- このうち、両耳間差情報と左右方向知覚の関係については多くの定量的な研究があるが、前後・上下の方向知覚を決定する具体的なスペクトル情報については、明らかにされていなかった。

著者らは、HRTF を複数のスペクトラル・ピークとスペクトラル・ノッチに分解し (Fig.2), その一部または全部で再構成した parametric HRTF を用いて音像定位実験を行い、以下のことを明らかにした[3].

- 1) 全ての Peak/Notch で再構成した parametric HRTF は実測 HRTF と同等の定位精度を有する。
- 2) ある特定の Peak/Notch は上昇角知覚に重要な役割を果たし、ある Peak/Notch は果たさない。
- 3) N1,N2 および P1 で再構成した parametric HRTF は実測 HRTF と同等の定位精度を有する (Fig.3,4).

また、正中面内の Peak/Notch の分布を観察すると、N1,N2 の周波数は音源の上昇角に依存して顕著に変化するが、P1 は変化しないことがわかった (Fig.5).

これらの結果より、N1 と N2 は上昇角知覚のスペクトラルキューであり、P1 は聴覚が N1 と N2 を分析する上でリファレンス情報として貢献していると考えられる。

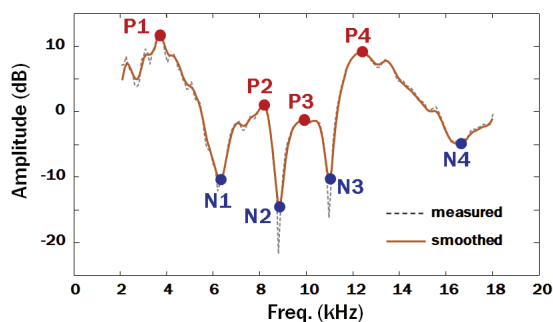


Fig.2 Examples of extracted spectral peaks and notches from measured HRTF.

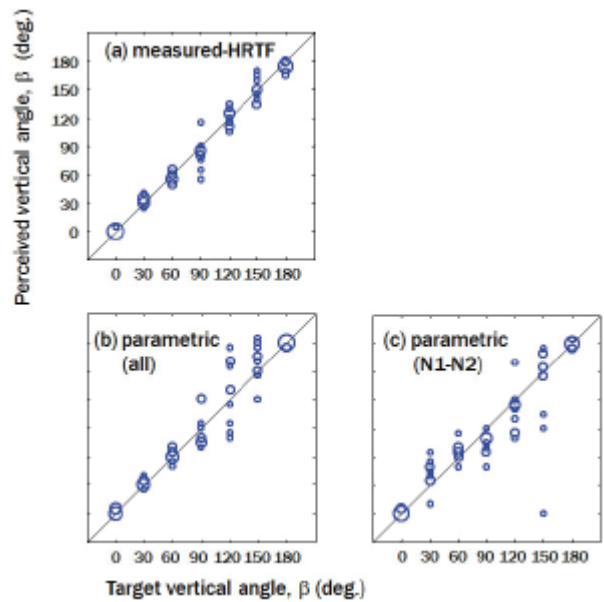


Fig.3 Responses of subject IT to stimuli of measured HRTFs and parametric HRTFs in the median plane

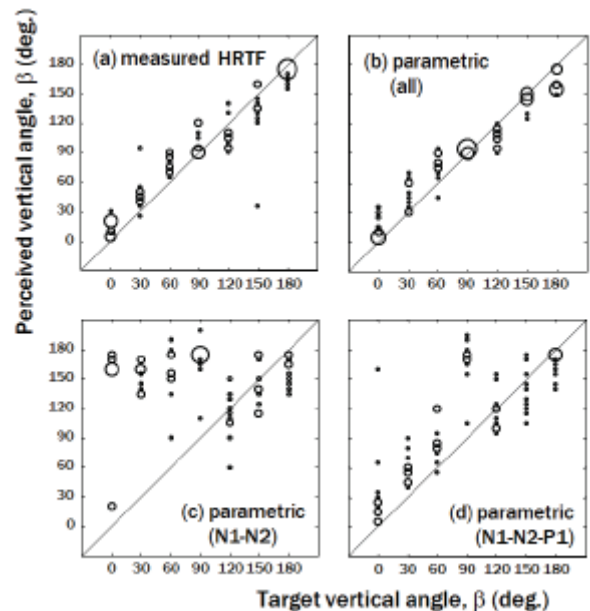


Fig.4 Responses of subject MK to stimuli of measured HRTFs and parametric HRTFs in the median plane.

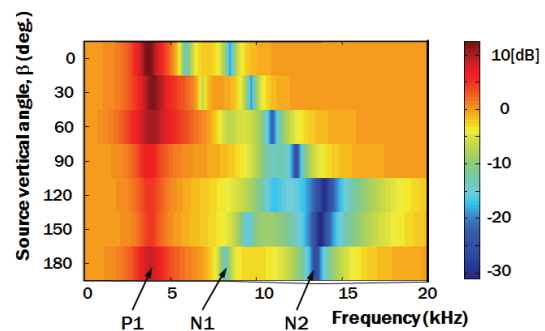


Fig.5 Distribution of frequencies of N1, N2, and P1 in the upper median plane

3 個人の方向知覚の cue の推定方法

方向知覚の cue は両耳間差および N1,N2 周波数であることがわかったので、次に個人のそれらの値を求めることを考えてみる。受聴者ひとりひとりについて、音源方向ごとにこれらの値を測定、あるいは数値計算で求めることは事実上不可能であり、これらの値を簡便に推定する方法が必要となる。

3.1 両耳間差の推定

両耳間差は頭部寸法（頭幅，頭長）との相関が高いことが知られており[4]，比較的容易に求めることができる。しかも，両耳間差は多少の誤差は問題にならないと考えられる。例えば，本人の HRTF を使った場合に側方角 30° に定位していたのが，他人の HRTF を用いたことによって 32° になったとしても，実用上大きな問題はないであろう。

3.2 N1,N2 周波数の推定

N1,N2 周波数については，推定誤差が大きいと，前後・上下誤判定を引き起こすため，高い推定精度が求められる。N1,N2 周波数は耳介部位，特に concha の寸法の影響が強いことが報告されているが[5]，定量的関係については未解明のままである。この問題については，数値計算による感度解析が有効であると考えられる。著者と NICT/ATR の研究グループでは，FDTD を用いて耳介形状と N1,N2 周波数の関係について検討を進めている[6]。この研究はまだ緒についたばかりであるが，今後，個人の耳介形状から N1,N2 周波数を求める予測式を導きたい。

4 HRTF の個人差の物理評価指標

いま，推定した受聴者個人の cue を基にして，多数の HRTF が蓄えられたデータベースから，その受聴者に適合した HRTF を選択することを考えてみる。この戦略を実現するには，cue の個人差を表現する指標（物差し）を作り，その指標における個人差の許容範囲（目盛り）を明らかにする必要がある。

4.1 従来の評価指標

従来，HRTF 間の類似度，あるいは適合度を評価する際に，スペクトル情報については

(1)式に示す SD (Spectral Distortion)が便宜的に使われてきた[7]。しかし，前後・上下方向の知覚の cue は N1,N2 周波数であり，全ての周波数成分を等しい重みで扱う SD で評価できるとは考えにくい。

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

4.2 Notch Frequency Distance (NFD)

まず，N1,N2 周波数の個人差を観察してみる。Fig.6 は男女 51 人の正面方向の HRTF から抽出した N1,N2 の周波数の分布である。N1 は約 6.2kHz から 9.6kHz，N2 は約 8.2kHz から 12.5kHz の範囲に広く分布しており，HRTF の個人差がよく表われている。また，概ね左下から右上に布置しており，N1 周波数が高い受聴者は N2 周波数も高くなる傾向がある。

ここで，このような HRTF 間のノッチ周波数の差に着目した Notch Frequency Distance (NFD)を提案し，HRTF の類似度の評価を試みる。HRTF_j と HRTF_k に関する NFD を式(2), (3)のように定義する。NFD₁ は N1 の周波数距離を，NFD₂ は N2 の周波数距離を表す。

$$NFD1_{j,k} = \log_2 \{ f_{N1}(HRTF_j) / f_{N1}(HRTF_k) \} \quad [\text{oct}] \quad (2)$$

$$NFD2_{j,k} = \log_2 \{ f_{N2}(HRTF_j) / f_{N2}(HRTF_k) \} \quad [\text{oct}] \quad (3)$$

ここで， f_{N1} ， f_{N2} はそれぞれ N1,N2 の周波数を表す。

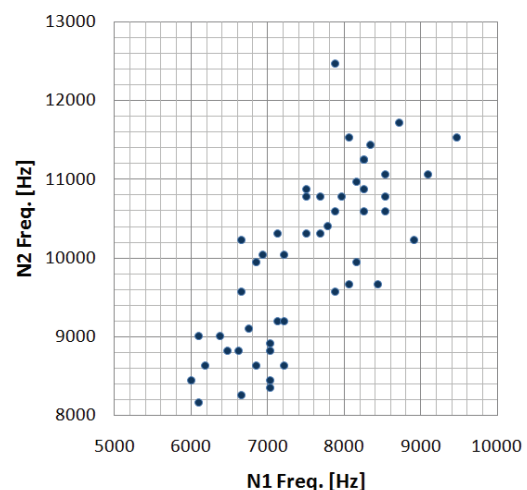


Fig.6 Distribution of N1 and N2 frequencies of 51 subjects for vertical angle of 0 degree in the median plane.

4.3 NFD による被験者間の HRTF の評価

NFD による HRTF 個人差評価の有効性を検証するため、51 人の正面方向の HRTF を用いて無響室で音像定位実験を行った。実験システムは、ノート PC(Dell 1525), オーディオインタフェース (RME Hammerfall DSP), イヤースピーカ (AKG K1000), およびイヤーマイク [3]より構成される。音源信号は 280Hz - 17kHz のホワイトノイズで、刺激の提示時間は 1.2 秒 (前後に 0.1 秒の立上がり/下がりを含む) である。被験者は無響室の中央にイヤースピーカとイヤーマイクを装着した状態で座り、まず再生系の伝達関数 $C(\omega)$ を測定した。測定後はイヤーマイクを取外し、刺激 $P(\omega)$ を式(4)により算出し、被験者に提示した。

$$P(\omega) = S(\omega) \cdot H(\omega) / C(\omega), \quad (4)$$

ここで、 $S(\omega)$ は音源信号、 $H(\omega)$ は HRTF である。被験者は知覚した音像の上昇角をマッピング法により回答した。全刺激をランダムな順で並べたものを 1 試行とし、10 試行からなる実験を行った。ターゲット方向は、正面方向である。被験者は男子学生 3 名である。

Fig.7 に代表的な被験者の各 HRTF に対する平均的的定位誤差を示す。横軸は、NFD である。ここで、NFD は式(5)で定義した。

$$NFD_{j,k} = |NFD1_{j,k}| + |NFD2_{j,k}| \quad [\text{oct.}], \quad (5)$$

この図は、NFD が 0.1oct 程度以内であれば被験者本人の HRTF とほぼ同等の定位精度が得られることを示している。これは、上昇角知覚における $N1, N2$ 周波数の jnd [8] の値によく対応している。一方、NFD が大きくなると、平均定位誤差は増大していく。ただし、NFD が大きくても精度よく定位できる HRTF も少なからず存在している。この理由については今後の検討が必要である。

5 おわりに

本稿では、HRTF を用いた音像制御、音場再生を実用化するには、HRTF の個人化が必要不可欠であるとの認識に立ち、一般の受聴者を対象とした HRTF 個人化実現への展望を述べた。

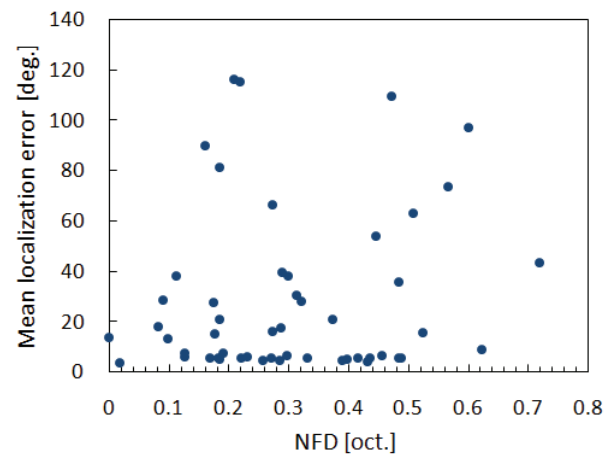


Fig.7 Mean localization error for others' HRTFs as a function of NFD.

音像制御、音場再生を実現するためには、HRTF の個人化に加えて、再生系伝達関数も個人に適應して補正する必要がある。ヘッドホン再生では、個人の外耳道内の伝達特性を予測して補正する必要があるが、萌芽的な研究が進みつつある[9]。さらなる研究の進展が期待される。

謝辞

本研究を進めるにあたり、伊藤元邦博士(パナソニック)に協力いただいた。また、HRTF の測定および Peak/Notch の抽出において板垣篤恵さん(神戸大大学院/現・竹中工務店)と東岡泰一君(神戸大大学院)の、音像定位実験において和田万正君(千葉工大)の協力を得た。記して感謝する。

参考文献

- [1] M. Morimoto and Y. Ando, J. Acoust. Soc. Jpn. (E), vol.1, pp.167-74, 1980.
- [2] M. Burkhard and R. Sachs, J. Acoust. Soc. Am., vol. 58, pp.214-222, 1975.
- [3] K. Iida *et al.*, Applied Acoustics, vol.68, pp.835-50, 2007.
- [4] K. Watanabe *et al.*, J. Acoust. Soc. Am., vol. 122, pp.2832-2841, 2007.
- [5] 板垣, 神戸大学修士論文, 2008
- [6] 竹本他, 音講論 (春), 1-9-2, 2009.
- [7] N. Inoue *et al.*, Acoust. Sci. & Tech. vol.26, pp.453-5, 2005.
- [8] 飯田他, 音講論 (春), pp.413-4, 2007.
- [9] 西村他, 信学技報 EA2008-104, pp.19-24, 2008.