

頭部伝達関数の4kHz以下の帯域の前後方向知覚の手掛かり*

◎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

1960年代以降, 正中面音像定位の手掛かりについて多くの研究が進められ, 頭部伝達関数の5kHz以上の帯域の振幅スペクトルが重要であることが明らかにされている。

しかし, 正中面音像定位のうち前後方向の判別においては, 4kHz以下の成分だけでも可能であることが示されている[1]. また水平面においても, 広帯域信号と比較すると若干精度は劣化するものの4kHz低域通過ノイズでも音像定位が可能である[2]. したがって, 4kHz以下にも何らかの前後方向知覚の手掛かりが存在すると考えられる。

本研究では4kHz以下の帯域における音像の前後判定の手掛かりを検討する。また, 得られた手掛かりをパラメトリックHRTFモデルに追加し, 音像定位精度の向上を検証する。

2. 4kHz以下の前方と後方のHRTFの比較

4kHz以下の前後方向知覚の手掛かりを探るため, 成人男女118耳介のHRTFを用いて後方と前方のスペクトルレベルの差を求めた(Fig. 1). 各ドットが各耳介の(後方-前方)のレベルを表している。ほとんどの周波数で0dB以下に布置されており, 後方より前方の音圧レベルの方が大きいと言えるが, 4kHz以上では個人差が大きい。一方で, 1kHz付近では後方の音圧レベルの方が大きく, 個人差も比較的小さい(青の矢印)。

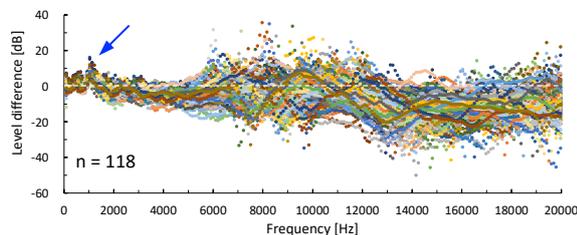


Fig. 1 Amplitude level difference of the HRTFs in the rear-front direction for 118 ears.

そこで, 1500Hz以下の帯域において, 後方と前方のHRTFのレベル差の118耳介の平均値と95%信頼区間を算出した(Fig. 2).

平均レベル差は1031.25Hzで最大となり, 5.6dBであった。95%信頼区間は0.48dBであった。

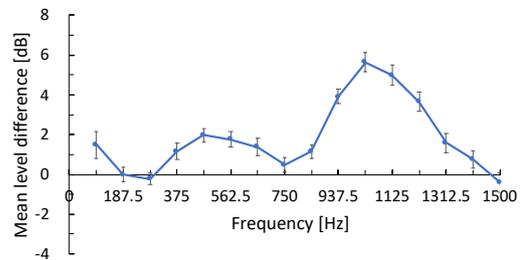


Fig. 2 Mean amplitude level difference of the HRTFs in the rear-front direction for 118 ears. Bars denote the 95% significance range.

平均レベル差が最大となる周波数を求めると, 全耳介の49%が1031.25Hzでレベル差が最大となった(周波数分解能: 93.75Hz). また, 86%の耳介が1031.25Hzを中心とする1/3 oct.の帯域でレベル差が最大となった。

このように, 後方のHRTFのレベルは前方と比較して1kHz付近で卓越し, その個人差も小さいことが明らかになった。

3. 1kHzの卓越周波数帯域と方向決定帯域との関係

1kHzの卓越周波数帯域が前後方向知覚の手掛かりとして寄与するか否かを検討するため, 方向決定帯域と比較する。

Blauertは正中面において音源方向に関わらず特定の方向(前方, 上方, 後方)に音像を知覚する帯域が存在することを発見し, 方向決定帯域(Directional Band)と呼んだ[3].

後方と前方のHRTFの振幅レベル差の118耳介の平均値と方向決定帯域を比較した結果, 前方および後方の卓越周波数帯域はそれぞれの方向決定帯域と一致した。しかし, 方向決定帯域は広帯域信号に対しては方向決定帯域だけを強調してもその方向に融合した音像を知覚しないことが報告されている[4-6].

そこで, 方向決定帯域が広帯域信号に対しても方向知覚の手掛かりとして寄与するか否かについて, 次章で検討する。

* A cue for front-back sound image localization in the head-related transfer function below 4kHz, by NAKAMURA, Fuka and IIDA, Kazuhiro (Chiba Institute of Technology).

4. 方向決定帯域は広帯域信号の方向知覚の手掛かりとなり得るか？

本章では 1kHz の卓越周波数帯域が広帯域信号に対しても後方知覚の手掛かりとして寄与するか否かを検討する。

4.1 広帯域信号の上昇角知覚の手掛かり

われわれは実測 HRTF の振幅スペクトルを複数のノッチとピークに分解し、その一部または全部を用いて再合成したパラメトリック HRTF (以降 pHRTF) を提案し、上半球正中面のいずれの方向においても実測 HRTF と同等の音像定位精度を提供する最小の pHRTF 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) (以降 N/P と記述) であることを見出した [4, 6]。

Fig. 3 に正中面の HRTF を示す。N1, N2 とともに音源が前方から上昇するにつれて周波数が高くなり、120° 付近で最も高くなる。このような音源方向依存性が上昇角知覚の手掛かりとして機能していると考えられる。一方、P1, P2 周波数は音源方向に関わらずほぼ一定である。音源方向に依存しない P1, P2 が上昇角知覚に必要である理由については諸説あるが [7]、定説となるものはない。

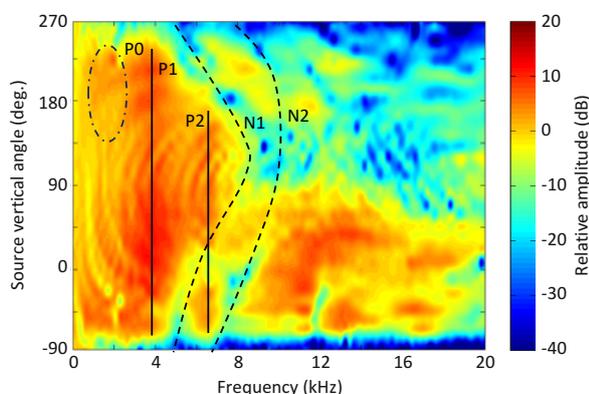


Fig. 3 Relationship between vertical angle of a sound source and frequencies of N1, N2, P1, P2, and P0 in the HRTFs in the median plane.

4.2 方向決定帯域と P1・P2・P0 の関係

方向決定帯域と P1, P2 周波数を比較する。Fig. 4 の橙色の実線は後方と前方のレベル差の 118 耳介の平均値、橙色の領域と緑色の領域はそれぞれ後方および前方の方向決定帯域を示す。緑色の破線は 118 耳介の P1 周波数の平均値である。P1 周波数は前方の方向決定帯域と一致していることがわかる。

同様に、青色の実線は上方と前方のレベル差の平均値、青色の領域は上方の方向決定帯

域を示す。青色の破線は 118 耳介の P2 周波数の平均値である。P2 周波数は上方の方向決定帯域と一致している。

ここで、Fig. 3 をもう 1 度観察すると、P1 では前方において、P2 では上方において、それぞれ赤色が濃い (ピークレベルが大きい) ことが示されている。P1 と P2 は音源方向による周波数の変化はないが、前方と上方の特徴がレベルの大小に表れ、それがそれぞれ前方および上方知覚の手掛かりとして寄与していると解釈することができる。

さらに Fig. 3 の 1kHz 付近を詳細にみると、前方や上方では 0dB 以下であるが 180° 付近では 5dB 程度卓越していることが観察される (1 点鎖線の楕円)。P1, P2 と同様に、1kHz の卓越周波数帯域に対応するピーク (以降 P0 と呼ぶ) は後方知覚の手掛かりとして寄与していると考えられる。Fig. 4 の橙色の波線はこの P0 を示す。P0 は後方の方向決定帯域 (橙色の領域) と一致している。

このように P0, P1, P2 は後方、前方、上方の方向決定帯域と一致すると言える。

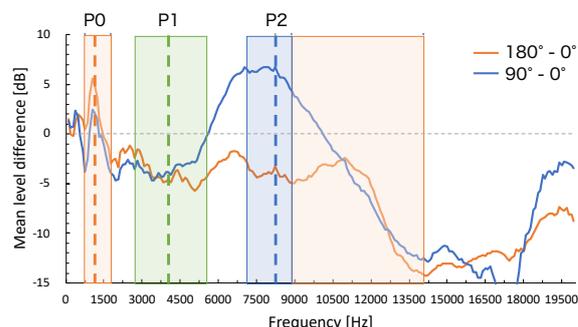


Fig. 4 Relationship between the directional bands for the rear (orange), the front (green), and above (blue) and the P0, P1, and P2. Solid orange and blue lines denote the mean amplitude difference of the HRTFs in the rear-front direction and above-front direction, respectively.

4.3 方向決定帯域と N/P を統合した仮説

以上より「上昇角を精密に知覚する手掛かりは HRTF のノッチであるが、ピークはノッチとともに存在することで後方、前方、上方の判別の手掛かりとして寄与する」という仮説を導くことができる。これは狭帯域信号の方向決定帯域による聴覚現象と、広帯域信号のノッチ・ピークによる聴覚現象の両方を矛盾なく説明することができる。

この仮説が妥当であるとすれば、1kHz の卓越周波数帯域 (P0) は、ノッチとともに存在することで広帯域信号に対して後方知覚の手掛かりとして寄与すると考えられる。

5. 1kHz の卓越周波数帯域が後方知覚に及ぼす影響

音響心理実験により 1kHz の卓越周波数帯域が後方知覚に及ぼす影響を検証した。

5.1 pHRTF(N1N2P1P2+P0)の作成

N1,N2,P1,P2 で再構成した従来の pHRTF モデルに P0 を加え, pHRTF(N1N2P1P2+P0) を作成した. P0 は他の N/P と同様に 2 次 IIR のピーキングフィルタで作成した.

ある被験者の後方の実測 HRTF と pHRTF を Fig. 5 に示す. P0 を加えることによって 1kHz 付近で実測 HRTF のスペクトルに近づくことが観察される.

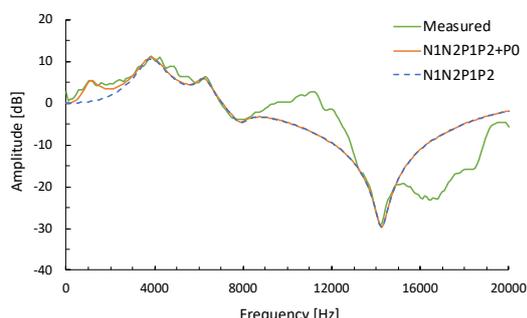


Fig. 5 Comparison among a measured HRTF and parametric HRTFs (N1N2P1P2 and N1N2P1P2 + P0) for the rear direction.

5.2 音像定位

5.2.1 実験方法

HRTF は以下の 4 種類を用いた. (1) 0°(前方)の被験者本人の実測 HRTF, (2) 180°(後方)の被験者本人の実測 HRTF, (3) 180°の被験者本人の N1N2P1P2 で構成した pHRTF, (4) 180°の被験者本人の N1N2P1P2+P0 で構成した pHRTF. 0°の実測 HRTF は実験システムと被験者の定位能力の妥当性を確認するために含めた.

音源信号は広帯域白色雑音 (200Hz-17kHz) である. 音源信号に HRIR を畳み込んだ刺激を FEC (Free air Equivalent Coupling to the ear) ヘッドホンとみなせる K1000 (AKG) により被験者に提示した. 外耳道入口を解放した状態での音圧を再現するため, ヘッドホン伝達関数の補正フィルタ処理を施した[8]. 提示音圧レベルは被験者の左右の外耳道入口で 63dB SPL である. 提示時間は 1.2s(立上り, 立下り各 0.1s を含める)である. 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上昇角を回答した.

5.2.2 実験結果

A 回答分布

各被験者の回答を Fig. 6 に示す. 0°と 180°の実測 HRTF では, いずれの被験者も概ね目標方向に回答した. 180°の N1N2P1P2 では, 被験者 1,2 は音像がやや上昇した. 被験者 3 は目標方向付近に回答した. N1N2P1P2+P0 では, 被験者 1,2 の回答は N1N2P1P2 の回答よりも目標方向に近い位置に分布した. 被験者 3 の回答は目標方向付近に分布した.

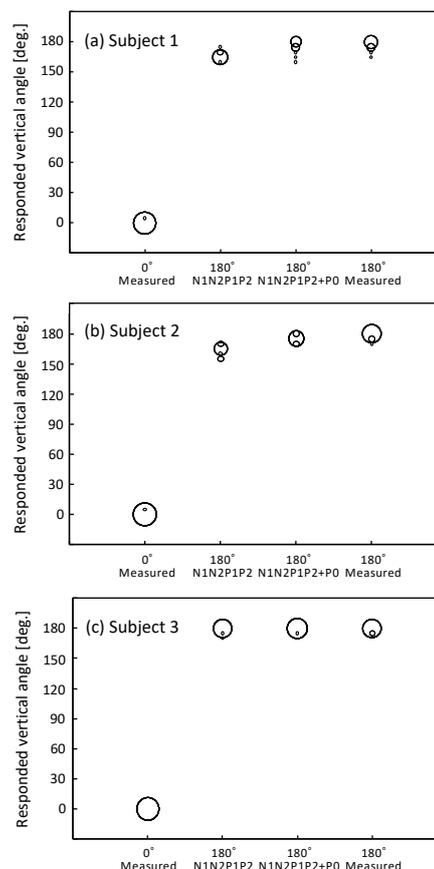


Fig. 6 Responses to the measured HRTF, N1N2P1P2, and N1N2P1P2+P0 for the front and rear directions.

B 平均上昇角誤差

平均上昇角誤差を Table 1 に示す. 0°と 180°の実測 HRTF では, 被験者の平均値は 1.0°と 2.5°であった. 180°の N1N2P1P2 では 10.9°, N1N2P1P2+P0 では 4.4°であった. いずれの被験者においても N1N2P1P2 に P0 を付加することにより上昇角誤差は低下した.

Table 2 に目標方向が 180°の実測 HRTF, N1N2P1P2, N1N2P1P2+P0 の Tukey の多重比較検定の結果を示す. 実測 HRTF と N1N2P1P2 の間には 3 名中 2 名で有意な差

が認められた($p < 0.01$). 一方, 実測 HRTF と N1N2P1P2+P0 の間にはいずれの被験者も有意な差は認められなかった. N1N2P1P2 と N1N2P1P2+P0 の間には 2 名で有意な差が認められた.

Table 1 Mean vertical angle error [deg.].

Subject	0°	180°	180°	180°
	Measured	N1N2P1P2	N1N2P1P2+P0	Measured
1	1.3	13.5	6.6	4.3
2	1.2	17.1	5.1	2.0
3	0.4	2.0	1.4	1.3
Ave.	1.0	10.9	4.4	2.5

Table 2 Results of Tukey's multiple comparison test among the measured HRTF and parametric HRTFs for the rear direction. **: $p < 0.01$, *: $p < 0.05$.

Subject	Measured	N1N2P1P2	N1N2P1P2+P0
1	Measured	-	-
	N1N2P1P2	**	-
	N1N2P1P2+P0	-	*
2	Measured	-	-
	N1N2P1P2	**	-
	N1N2P1P2+P0	-	**
3	Measured	-	-
	N1N2P1P2	-	-
	N1N2P1P2+P0	-	-

5.3 みかけの音源の幅

シェッフェの対比較の浦の変法を用いてみかけの音源の幅(ASW: Auditory Source Width)を比較する実験を行った.

5.3.1 実験方法

実験は無響室で行った. HRTF は以下の 3 種類を用いた. (1)被験者本人の後方の実測 HRTF, (2)被験者本人の後方の N1N2P1P2, (3)被験者本人の後方の N1N2P1P2+P0. 音源信号は広帯域白色雑音(200Hz-17kHz)で, 提示時間は 1.0s である. 音源信号に HRIR を畳み込んだ刺激(3 種類)を FEC ヘッドホンとみなせる 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. 提示音圧レベルは 68dB である. 被験者は正常な聴力を持つ成人女性 2 名である.

2 つの刺激を対にして提示し, 先に提示された刺激に対して, 後に提示された刺激の ASW が狭いか広いかを回答させた. 非常に狭い場合は -2.0, どちらでもない場合は 0.0, 非常に広い場合は 2.0 とした. 被験者は, -2.0 から 2.0 までの数直線に 0.1 間隔の目盛をつけた回答用紙にペンで印をつけて回答した.

刺激対は $3 \times 2 = 6$ 種類であり, 各刺激対をランダムな順に 5 回ずつ被験者に提示した.

5.3.2 実験結果

各刺激の布置を Fig. 7 に示す. ASW が最も広いのは N1N2P1P2 で, N1N2P1P2+P0

と実測の ASW は狭い.

有意水準 5%のヤードスティックと比較すると, 実測と N1N2P1P2+P0 には有意差は認められなかった. 一方, N1N2P1P2 は, 実測とも N1N2P1P2+P0 とも有意な差が認められた.

つまり, N1N2P1P2 の音像の空間的な広がりには実測 HRTF よりも広がる(像がぼやける)が, P0 を付加することによって実測 HRTF と同等の広がりになると言える.

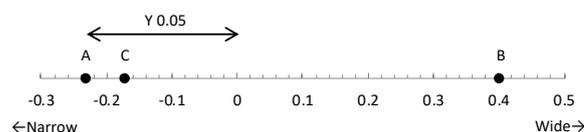


Fig. 7 ASW for the measured HRTF (A), N1N2P1P2 (B), and N1N2P1P2+P0(C) for the rear direction.

6. 考察

1kHz 付近で後方の HRTF が前方の HRTF と比較して卓越する理由はなんだろうか. 波長との関係から耳介の影響は考えにくい. また, 個人差が小さいことを併せて考えると, 胴体(torso)が影響している可能性がある. そこで, 大谷ら[9]が胴体有り無しの 2 条件で測定した 4128C HATS (B&K)の HRTF を見比べたが, この測定結果からは胴体の存在と 1kHz 付近での後方の HRTF の卓越との間に関連は見出せなかった. 卓越周波数帯域の成因については今後さらに検討する必要がある.

7. おわりに

本研究では, 4kHz 以下の帯域の音像の前後方向知覚の手掛かりを検討した. その結果, 後方の 1kHz の卓越周波数帯域が寄与していることが示唆された.

参考文献

- [1] 森本と斉藤, 日本音響学会聴覚研究会資料, H-40-1 (1977)
- [2] 森川と平原, 日本音響学会誌 68 巻 41, pp.71-179 (2012)
- [3] Blauert, ACUSTICA, 22, pp.205-213, 1969
- [4] Iida *et al.*, Appl. Acoust. 68, 835-850, 2007.
- [5] 竹内と飯田, 日本音響学会講演論文集, pp.665-668(2017年9月)
- [6] Iida *et al.*, Appl. Acoust. 129, 239-247, 2018.
- [7] Iida, Head-Related Transfer Function and Acoustic Virtual Reality, pp.36-38, Springer (2019)
- [8] Møller *et al.*, J Audio Eng. Soc., 43, pp.203-17 (1995)
- [9] 大谷他, 日本音響学会講演論文集, pp. 515-516(2008年3月)