

頭部伝達関数の第2ピークが正中面上方の音像定位に及ぼす影響

飯田 一博[†] 石井 要次[‡]

{[†] 先進工学部, [‡] 大学院工学研究科 } 千葉工業大学

275-0016 千葉県習志野市津田沼 2-17-1

E-mail: [†] kazuhiko.iida@it-chiba.ac.jp, [‡] s1179501rq@s.chibakoudai.jp

あらまし 4kHz 付近の最も周波数の低いピーク(P1)と P1 より高い周波数で最も周波数の低い 2 つのノッチ(N1, N2)だけで再構成されたパラメトリック頭部伝達関数(HRTF)は, 前方と後方において実測 HRTF と同等の音像定位精度を有することが報告されている. しかし, 上方については, 被験者によっては音像定位精度が低下する. 本論文では, 上半球正中面において 2 つの音像定位実験を行った. 実験結果は以下のとおりである. (1) N1N2P1 に P2 を加えることにより, 90°および 120°で音像定位精度が向上した. (2) N1N2P1+P2 と実測 HRTF の音像定位誤差の差は, 目標とする 7 方向すべてにおいて 10°未満であった. (3) P2 だけを再現しても音像は上方に知覚されなかった. また, HRTF を観察したところ, 90°での N1 のレベルはノッチの検知限を超えないが, P2 を加えることにより N1 の相対レベルは検知可能となることが明らかになった. 以上より, P2 それ自身はスペクトラルキューではないが, P2 は N1 を強調することにより, 正中面上方の音像定位に重要な役割を果たしていると言える.

Keyword 頭部伝達関数, 音像定位, 正中面, スペクトラルキュー, ピーク, ノッチ

Effects of the second lowest frequency spectral peak in head-related transfer functions on the sound localization for the upper direction in the median plane

Kazuhiro IIDA[†] and Yohji ISHII[‡]

{[†] Faculty of Advanced Engineering, [‡] Graduate School } Chiba Institute of Technology

2-17-1 Tsudanuma, Narashino, Chiba, 275-0016 Japan

E-mail: [†] kazuhiko.iida@it-chiba.ac.jp, [‡] s1179501rq@s.chibakoudai.jp

Abstract It is reported that the parametric head-related transfer function (HRTF) recomposed of only the lowest frequency peak around 4 kHz (P1) and the two lowest frequency notches (N1 and N2) above P1 provided approximately the same localization performance as the measured HRTFs for the front and rear directions. However, for the upper directions, the localization performance for some of the subjects decreased. In the present study, two localization tests were carried out in the upper median plane. The results revealed the following: (1) the localization performance was improved by adding P2 to N1N2P1 at the vertical angles of 90° and 120°; (2) the difference in localization error between N1N2P1+P2 and the measured HRTF was less than 10° at all seven target vertical angles; and (3) a sound image was not perceived in the upward direction by reproducing only P2. An observation of the HRTFs revealed that the level of N1 doesn't exceed the threshold for the notch detection at a vertical angle of 90°. However, by adding P2, the relative level of N1 becomes detectable. Therefore, P2 is not a spectral cue, but, by enhancing N1, plays an important role in the localization for the upper directions in the median plane.

Keyword head-related transfer function, sound localization, the median plane, spectral cue, peak, notch

1. はじめに

ヒトの頭部伝達関数(HRTF: head-related transfer function)の 5 kHz 以上のノッチやピークが音像の上昇角知覚に顕著に貢献していることは広く知られている. Shaw and Teranishi (1968)は音源の仰角が-45°から 45°に変化するとノッチ周波数が 6 kHz から 10 kHz に変化することを報告している. Butler and Belendiuk (1977)

は周波数応答曲線における明確なノッチは, 音源が前半球正中面の上方から下方に移動すると, 低い周波数に移動することを示した. Raykar *et al.* (2005)は耳介による深いノッチは仰角知覚に重要な特徴であり, ノッチ周波数は耳介の物理的な寸法や形状に関連すると報告している.

ノッチやピークの生成メカニズムも研究されてい

る。Gardner and Gardner (1973)は耳介の窪みを順次埋めて正中面音像定位実験を行い、耳介の閉塞を進めると音像定位性能が低下すること、および低域成分と比較して高域成分がより重要であることを示した。Musicant and Butler (1984)は、4 kHz 低域通過ノイズと比較して、4 kHz 高域通過ノイズは有意に重要であること、および4 kHz 高域通過ノイズにおいては耳介の閉塞により多くの前後誤判定が発生することを報告した。Iida *et al.* (1998)は耳介の主要な3つの窪み、すなわち舟状窩(scapha)、三角窩(fossa)、耳甲介腔(concha)を粘土で閉塞して音像定位実験とHRTF測定を行った(粘土には外耳道と同じ直径の穴を開け、外耳道を外と繋げている)。その結果、3 kHz 以下では全ての窪みを閉塞しても、そのHRTFは閉塞していない耳介のHRTFと同等であるが、耳甲介腔を閉塞するだけで4 kHz 付近のピークと4 kHz から10 kHz の間の鋭いノッチが消滅した。また、耳甲介腔を閉塞したHRTFの音像定位精度は閉塞していないHRTFと比較して有意に低下した。これらより、彼らは正中面音像定位のスペクトラルキューは耳甲介腔の伝達関数の4 kHz 以上に存在すると結論付けた。Takemoto *et al.* (2012)はFDTD(finite-difference time-domain)法を用いて頭部形状からHRTFを計算し、耳介がHRTFの基本的なピーク・ノッチパターンを決定していると報告した。ピークは耳介のノーマルモードと一致する。ノッチ周波数では耳甲介舟(cymba)と三角窩に腹が生じ、耳甲介腔に節が生じる。

ノッチとピークについては、微細構造よりむしろ概形が重要であることが報告されている。Asano *et al.* (1990)はARMAモデルで平滑化したHRTFをヘッドホンで再生する正中面音像定位実験を行い、仰角判断の主要な手掛かりは5 kHz 以上の高域にあり、ピークやノッチの微細部分ではなく巨視的なパターンの情報が利用されていることを示した。Kulkarni and Colburn (1998)はチューブホン(tube-phones)を用いて、水平面内(方位角: -45° , 0° , 45° , 180°)の被験者本人のHRTFの振幅スペクトルを7段階で平滑化して音像定位実験を行い、微細なスペクトル構造は、ノッチやピークの概形と比較すると、相対的に重要ではないことを示した。ただし、ノッチの1つが消滅するような過度な平滑化を施すと、音像の仰角は上昇した。

上昇角知覚のスペクトラルキューについても研究が進められている。Hebrank and Wright (1974)は上半球正中面(-30° から $+210^\circ$ まで、 30° 間隔)に設置されたスピーカを用いて無響室で音像定位実験を行った。刺激は帯域幅が low-pass, high-pass, band-pass, band-stop filters で制限された白色雑音である。彼らは16 kHz 以上もしくは3.8 kHz 以下のスペクトルエネルギーは音

像定位性能に影響しないこと、つまり正中面定位に利用されている特性は3.8から16 kHz であることを示した。さらに彼らは、フィルタ処理なしと比較して、ある方向に1.5倍以上の回答が得られたフィルタをスペクトラルキューと定義し、前方のキューは低域側のカットオフ周波数が4 kHz から8 kHz の間にある1オクターブノッチと13 kHz 以上のエネルギーの上昇、上方のキューは7 kHz から9 kHz の間の1/4オクターブのピーク、そして後方のキューは10 kHz から12 kHz の間にある小さなピーク(このピークの高域側および低域側でのエネルギーの減衰を伴う)であると報告した。

Iida *et al.* (2007)は上昇角音像定位をシミュレートするパラメトリックHRTFモデルを提案した。パラメトリックHRTFは、音源の上昇角に依存しない4 kHz のピーク(Shaw and Teranishi, 1968)を下限周波数として、受聴者の実測HRTFから抽出したノッチやピークから再構成される。ノッチとピークは周波数の順にラベルがつけられる(P1, N1, P2, N2, ...のように)。ノッチとピークは中心周波数、レベル、先鋭度でパラメトリックに表現される。彼らは上半球正中面において音像定位実験を行い以下のことを示した。(1) すべてのノッチとピークで再構成されたパラメトリックHRTFは被験者本人の実測HRTFと同等の定位性能を有する。(2) 4 kHz 付近の第1ピーク(P1)と、P1周波数より高く最も周波数の低い2つのノッチ(N1, N2)だけで再構成されたパラメトリックHRTFは、前方と後方においては被験者本人の実測HRTFと同等の定位性能を有する。(3) しかし上方においては、N1, N2, P1で再構成されたパラメトリックHRTFの音像定位性能は、被験者によっては本人の実測HRTFと比較して低下する。(4) N1とN2の周波数は音源の上昇角に強く依存するが、P1周波数はほぼ一定で上昇角に依存しない。これらの結果より、彼らはN1とN2は、少なくとも前方と後方では、スペクトラルキューとして重要な役割を果たしていると結論付けた。さらに彼らは、ヒトの聴覚システムが耳入力信号からN1とN2を分析するための参照情報としてP1を用いている可能性を指摘した。

先に述べたように、Hebrank and Wright (1974)は上方のスペクトラルキューは7 kHz から9 kHz の間の1/4オクターブのピークであると主張した。このピークはP2と一致し、さらにBlauert (1969/70)が提案した上方の方向決定帯域とも一致する。しかし、狭帯域信号を用いて得られた方向決定帯域が広帯域信号に対するスペクトラルキューの役割を果たすか否かは不明である。

本論文は2つの目的を有する。1つは、N1, N2, P1にP2を加え、N1, N2, P1, P2で再構成されたパラメトリックHRTFが上方において実測HRTFと同等の音像定位

性能を有するか否かを検証することである。もう一つの目的は、正中面上方の音像定位における P2 の役割を解明することである。

本論文では、方位角と仰角ではなく、側方角と上昇角で表現した耳軸座標系(Morimoto and Aokata, 1984)を用いる。上昇角は、音源を通る矢状面において、音源と耳軸を結ぶ直線と水平線がなす角と定義する。

2. パラメトリック HRTF の再構成方法

2.1. HRTF の測定

無響室において、22-24 歳の 3 人の男性被験者(MKI, OIS, OTK)および 1 人の女性被験者(YSD)の上半球正中面 7 方向(30°間隔)の HRTF を測定した。測定用信号は swept sine wave で、サンプリング周波数は 48 kHz である。測定用信号は上半球正中面に 30°間隔で設置した直径 80mm のスピーカ(FOSTEX FE83E)から提示した。スピーカから被験者の頭部中心までの距離は 1.2m である。被験者の外耳道入口で測定用信号を収録するため、耳栓型マイクロホン(Iida *et al.*, 2014)を用いた。マイクロホンの振動板は被験者の外耳道入口に位置し、blocked-entrances condition (Shaw and Teranishi, 1968)とみなせる。HRTF は以下のように求めた。

$$HRTF_{l,r}(\omega) = G_{l,r}(\omega)/F(\omega) \quad (1)$$

ここで、 $F(\omega)$ は被験者のいない自由空間において被験者の頭部中心に相当する位置で測定したインパルス応答 $f(t)$ のフーリエ変換であり、 $G_{l,r}(\omega)$ は自由空間において被験者の外耳道入口で耳栓型マイクロホンにより測定したインパルス応答のフーリエ変換である。

2.2. ノッチとピークの抽出

N1, N2, P1, P2 は耳介で生成されているため(Shaw and Teranishi, 1968; Lopez-Poveda and Meddis, 1996; Takemoto *et al.*, 2012), 0°と 180°については被験者の頭部インパルス応答(HRIR)の初期部分から抽出した。そのアルゴリズム(Iida *et al.*, 2014) を以下に示す。

- (1) HRIR の絶対値が最大となるサンプルを検出する。
- (2) 検出したサンプルを中心として 4 次 96 サンプルの Blackman-Harris 窓で HRIR を切り出す。
- (3) 全ての要素をゼロとした 512 ポイントの配列を用意し、(2) で切り出した HRIR を上書きする。ただし、最大サンプルを配列の 257 番目に合わせる。
- (4) FFT より 512 ポイントの配列の振幅スペクトルを求め、差分法で極大値および極小値をすべて検出する。

(5) 3 kHz 以上で最も周波数の低い 2 つの極大値をそれぞれ P1,P2 とし、P1 より高い帯域で最も周波数の低い 2 つの極小値を N1,N2 とする。

30°から 150°では、被験者によっては N1,N2 は浅くて不明瞭な場合があるので、以下のアルゴリズムにより N1,N2 を抽出した。

(1) FFT より 512 サンプルの HRIR の振幅スペクトルを求め、差分法で極小値をすべて検出する。

(2) Iida and Ishii (2011)が報告した以下の回帰式を用いて N1,N2 周波数を算出する。彼らは、正面方向の N1,N2 周波数は被験者に強く依存するが、上昇角の関数としての N1,N2 周波数の振舞いは被験者間で共通であるとみなせると報告している。

$$f_{N1}(\beta) = 1.001 \times 10^{-5} \times \beta^4 - 6.431 \times 10^{-3} \times \beta^3 + 8.686 \times 10^{-1} \times \beta^2 - 3.265 \times 10^{-1} \times \beta + f_{N1}(0) \quad [\text{Hz}] \quad (2)$$

$$f_{N2}(\beta) = 1.310 \times 10^{-5} \times \beta^4 - 5.154 \times 10^{-3} \times \beta^3 + 5.020 \times 10^{-1} \times \beta^2 + 2.563 \times 10 \times \beta + f_{N2}(0) \quad [\text{Hz}] \quad (3)$$

ここで f_{N1} と f_{N2} はそれぞれ N1,N2 周波数を表し、 β は上昇角である。

(3) (1)で検出した極小値のうち、(2)で算出した N1, N2 周波数から 0.2 オクターブ以内で最も深い極小値をそれぞれ N1,N2 とする。その理由は、上昇角知覚における N1,N2 の丁度可知差(JND)は 0.1 から 0.2 オクターブと考えられる(Iida and Ishii, 2011)からである。

P1,P2 周波数は上昇角に依存しないので、30°から 150°の P1,P2 は 0°と同様とした。

2.3. パラメトリック HRTF の生成

パラメトリック HRTF は、2 次の IIR フィルタで作成したノッチとピークを重ね合わせることで生成した。実測 HRTF, N1,N2,P1 で再構成したパラメトリック HRTF(以降 N1N2P1 と呼ぶ)、N1,N2,P1,P2 で再構成したパラメトリック HRTF(以降 N1N2P1+P2 と呼ぶ)の例を FIG. 1 に示す。パラメトリック HRTF は再構成したノッチとピークを精度よく再現し、その他の帯域のスペクトルは平坦である。

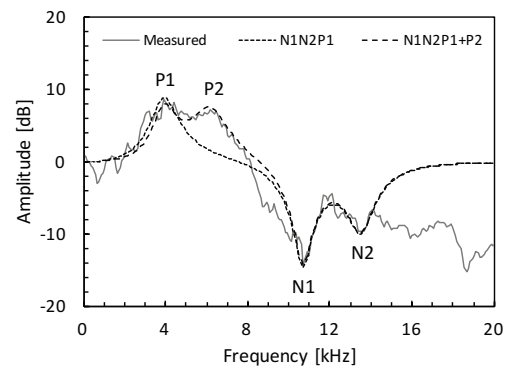


FIG. 1. Examples of a measured HRTF and parametric HRTFs (N1N2P1 and N1N2P1+P2) for a vertical angle of 90°.

3. 音像定位実験 1

N1N2P1+P2が上方において実測 HRTF と同等の音像定位精度を有するか否かを検証するため、上半球正中面において音像定位実験を行った。

3.1. 実験方法

被験者は 2.1 節で HRTF の測定を行った 4 人(MKI, OIS, OTK, YSD)である。HRTF を以下の 3 種類を用いた。(1) 被験者本人の実測 HRTF, (2) 被験者本人の N1N2P1, (3) 被験者本人の N1N2P1+P2。

音像定位実験は防音室で行った。室の内法は 4.6m (W) × 5.8m (D) × 2.8m (H)で、暗騒音は 19.5 dB(A)であった。実験システムの構成は、ノート PC (DELL XPS M1330), オーディオインタフェース(RME Fireface 400), アンプ (Marantz PM4001), A/D コンバータ (Roland M-10MX), オープンタイプヘッドホン, 耳栓型マイクロホンである。

外耳道入口を開放した状態における鼓膜での音圧 P は、閉塞した状態の外耳道入口の音圧に以下に示す補正フィルタ G を施してヘッドホンから提示することにより得られる(Møller *et al.*, 1995)。

$$P = S \times HRTF \times G, \quad (4)$$

$$G = \frac{1}{M \times PTF} \times \frac{Z_{ear\ canal} + Z_{headphone}}{Z_{ear\ canal} + Z_{radiation}} \quad (5)$$

$$\triangleq \frac{1}{M \times PTF} \times PDR, \quad (6)$$

ここで、 S は音源、 M は耳栓型マイクロホンの伝達関数、 $M \times PTF$ は閉塞した外耳道入口で測定したヘッドホンの伝達関数、 $Z_{ear\ canal}$ と $Z_{headphone}$ はそれぞれ外耳道およびヘッドホンのインピーダンス、 $Z_{radiation}$ は外耳道から外側をみた放射インピーダンスである。 G の第 2 項は PDR (pressure division ratio) と呼ばれ、Møller *et al.* は PDR が 1 となるヘッドホンを FEC (free-air equivalent coupling to the ear) ヘッドホンと定義した。本実験では、彼らが FEC ヘッドホンとみなせるとした 2 種類のヘッドホン、K1000 (AKG) および DT990 PRO (Beyerdynamic) (以降 DT990 と記述) を用いた。

K1000 では、 $M \times PTF$ の補正は 200 Hz から 17 kHz の範囲で行った。 $M \times PTF$ は以下の手順で測定した。被験者は防音室の中央に座り、耳栓型マイクロホンを外耳道に装着した。マイクロホンの振動板は 2.1 節の HRTF 測定と同様に外耳道入口に位置した。被験者にヘッドホンを装着して M 系列信号 (48kHz サンプリング, 12 次, 繰返しなし) を提示し、耳栓型マイクロホンまでの伝達関数 $M \times PTF$ を測定した。さらに、ヘッドホンを装着したまま、その位置を変えないようにして耳栓型マイクロホンだけを取り外した (耳介はヘッド

ホンに覆われていない (FIG. 2(a))). ヘッドホンから耳栓型マイクロホンまでの 200 Hz から 17 kHz までの代表的な peak-to-peak レンジは約 20 dB であったが、補正フィルタ G により約 3 dB まで低減した (FIG. 3)。

DT990 では、耳介がヘッドホンに覆われているため (FIG. 2(b)), $M \times PTF$ を測定するにはヘッドホンの着脱が必要となる。 $M \times PTF$ はヘッドホンの装着位置により変化するの、それを測定してもヘッドホンの着脱を行うと正確な補正は保証されない。以上の理由により、DT990 では $M \times PTF$ の補正は行わなかった。参考のため、FIG. 4 に 4 人の被験者の DT990 の $M \times PTF$ の測定例を示す。peak-to-peak レンジは、Møller *et al.* (1995) の報告と同様 30 dB を超える。

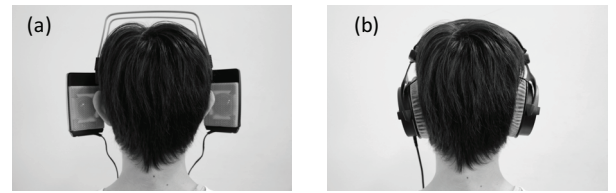


FIG. 2. Headphones used in the experiments. (a): K1000 (AKG), (b): DT990 (Beyerdynamic).

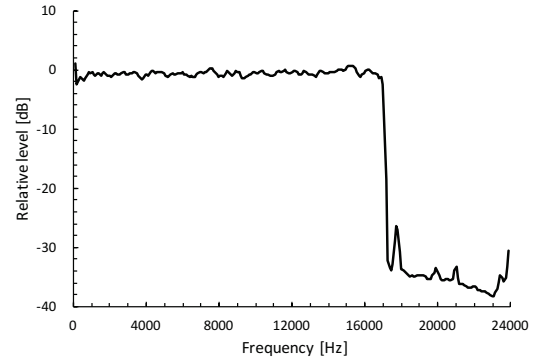


FIG. 3. An example of transfer functions between K1000 and the earplug-type microphone with a compensation filter, G .

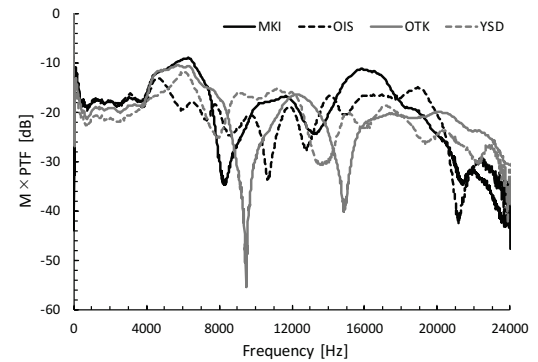


FIG. 4. Transfer functions between DT990 and the earplug-type microphone, $M \times PTF$.

K1000とDT990の再生条件の違いを表現するために、以降、K1000を用いて式(6)が1とみなせる条件をFEC条件、DT990を用いて $1/(M \times PTF)$ が残る条件を準FEC条件と表す。Møller *et al.* (1995)が定義したFECヘッドホンとの違いに注意されたい。準FEC条件は理論的厳密さに欠けるが、伝達関数の測定が不要であるため、一般の受聴者が使いやすいという利点を有している。つまり、準FEC条件は実用性の観点から検証した。

音源信号は200Hzから17kHzの広帯域ホワイトノイズである。刺激は被験者の左右の外耳道入口で63dB SPLで提示した。刺激の提示時間は1.2秒(立上がり、立下がりそれぞれ0.1秒を含む)である。目標方向は上半球正中面における30°間隔の7方向である。被験者が目標方向を推測することを避けるため、回答方法にはマッピング法を用いた。ノートPCの画面上に正中面を示す円周と矢印を表示した。被験者のタスクは、知覚した上昇角を円周上にスタイラスペンでクリックすることである。頭内定位した場合はチェックボックスに印を入れるよう指示した。

実測HRTFと2種類のパラメトリックHRTFの音像定位実験は別々に行った。Møller *et al.* (1995)は2つの条件、すなわち(1)1人の被験者のHRTFをランダム化した条件と(2)数人のHRTFをランダム化した条件を比較し、同じ刺激群を分離して実験しても混合して実験しても有意な差はないことを示している。

1回の試行で35刺激(7方向×5回)をランダム化して被験者に提示した。1試行の所要時間は約7分である。3種類のHRTFについて2試行ずつ実験を行った。つまり、被験者は各刺激に対して10回ずつ回答した。

3.2. 実験結果

3.2.1. FEC条件

A. 回答分布

FIG. 5に実測HRTF、N1N2P1、N1N2P1+P2に対する4人の被験者の回答を示す。縦軸は知覚した上昇角、横軸は目標上昇角である。円の直径は回答の頻度に比例し、回答方向は5°間隔で丸めている。

被験者MKIでは、実測HRTFに対する回答はほぼ対角線上に分布した。N1N2P1の回答は0°、30°、150°、180°では実測HRTFと同様であった。しかし、他の3方向(60°、90°、120°)では回答の分散は実測HRTFより大きい。N1N2P1+P2の回答は、0°、30°、150°、180°では実測HRTFおよびN1N2P1と同様であった。90°、120°では分散は実測HRTFよりやや大きく、N1N2P1よりやや小さい。しかし、60°では分散はN1N2P1より大きい。

被験者OISでは、実測HRTFに対する回答は逆S字カーブを描き、30°、60°、120°でやや上方寄りに回答する傾向があった。N1N2P1の回答は、0°、30°、60°、180°では実測HRTFと同様であったが、90°、120°、150°では

実測HRTFと比べてやや後方に分布する傾向があった。N1N2P1+P2の回答は、0°、30°、60°、180°では実測HRTFおよびN1N2P1と同様であった。90°、120°では回答分布はN1N2P1より実測HRTFに近づいた。しかし、150°では上方寄りに知覚した。

被験者OTKでは、実測HRTFに対しては0°、30°、90°、180°では目標方向に回答した。しかし、60°、120°、150°では90°付近に回答した。N1N2P1の回答は0°、150°、180°では実測HRTFと同様であった。しかし、30°では前方に寄る傾向がみられ、60°、90°、120°では0°と90°付近の両方に回答が分布した。N1N2P1+P2の回答は、すべての目標方向において実測HRTFと同様であった。N1N2P1で観測された回答の分離はみられなかった。

被験者YSDでは、実測HRTFに対する回答は逆S字カーブを描き、30°、60°、120°でやや上方に寄る傾向があった。N1N2P1の回答は、90°、120°、180°では実測HRTFと同様であった。しかし、0°、30°、60°では実測HRTFと比べてやや上方に分布する傾向があった。150°では目標方向付近に回答した。N1N2P1+P2の回答は60°を除きN1N2P1と同様であった。60°では回答の分散はN1N2P1より大きい。

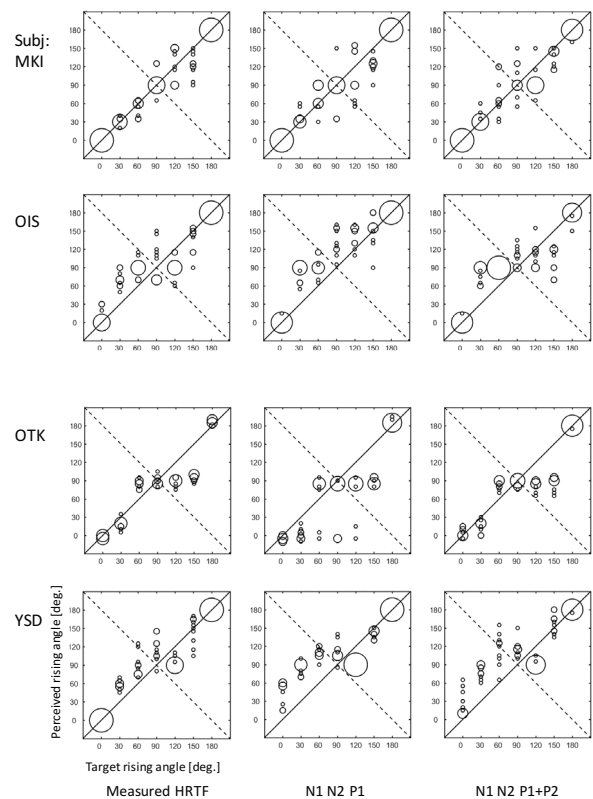


FIG. 5. Responses to the measured HRTF, N1N2P1, and N1N2P1+P2 under the FEC condition. The ordinate represents the perceived vertical angle, and the abscissa represents the target vertical angle. The diameter of each circle is proportional to the number of responses with a resolution of 5°.

B. 平均上昇角定位誤差

各 HRTF と各目標方向の平均上昇角定位誤差(E_v)を求めた(FIG. 6). ここで, 平均上昇角定位誤差は目標の上昇角と回答した上昇角の差の絶対値の平均と定義する. さらにパラメトリック HRTF と実測 HRTF の E_v の差 ΔE_v を求めた(FIG. 7).

実測 HRTF の E_v は, Carlile *et al.* (1997)や Majdak *et al.* (2010)の報告と同様, 水平面上の 0° , 180° では小さく(3.1° , 1.8°), 水平面から上昇した 30° から 150° では大きい(18.1 - 29.1°)傾向がある. N1N2P1, N1N2P1+P2 でも同様の角度依存性がみられた. N1N2P1, N1N2P1+P2 の E_v は実測 HRTF より大きい, N1N2P1+P2 の E_v は 0° , 30° , 90° , 120° で N1N2P1 より小さい. 特に 90° では 10.8° 減少した. ただし, 150° では N1N2P1 より増加した. これは被験者 OIS が上方寄りに回答したことに依るが, 上方に寄った原因は不明である.

N1N2P1 と実測 HRTF の E_v の差(ΔE_v)は 150° , 180° では小さく(0.6° , 0.4°), 他の上昇角においては 8.1° から 14.0° であった. N1N2P1+P2 の ΔE_v は 90° , 120° , 180° において小さく(-0.4° , 0.4° , 0.4°), それ以外の上昇角でも 5.9° から 8.9° であり, すべての上昇角で 10° 未満であった.

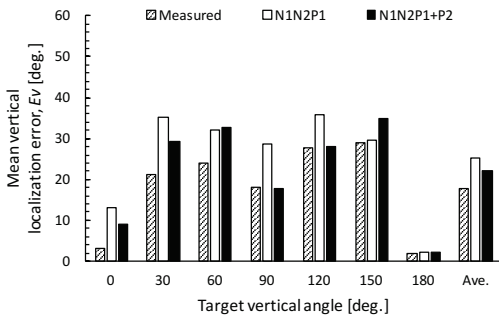


FIG. 6. Mean vertical localization error, E_v , for each HRTF and target vertical angle under the FEC condition.

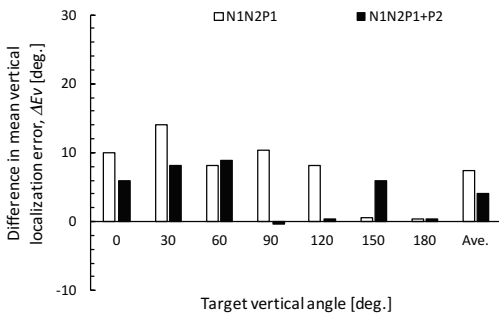


FIG. 7. Difference in E_v between the parametric HRTF and the measured HRTF, ΔE_v , the under FEC condition.

C. 前後誤判定率

TABLE I に前後誤判定率を示す. ここで, 前後誤判定率は上半球正中面において目標方向と異なる四分円に音像を知覚した率と定義する.

0° と 180° では, いずれの HRTF においても前後誤判定率は 0% であった. しかし, 他の 4 方向においては, N1N2P1 と N1N2P1+P2 の前後誤判定率は実測 HRTF のそれより高く, N1N2P1+P2 の前後誤判定率は N1N2P1 よりやや低い.

HRTF 間の前後誤判定率に統計的に有意な差があるかを明らかにするためにカイ 2 乗検定を行った. その結果, いずれの目標方向においても HRTF 間に有意な差は認められなかった.

TABLE I Ratio of front-back confusion for each HRTF and target vertical angle under the FEC condition.

HRTF	Target vertical angle ($^\circ$)						
	0	30	60	90	120	150	180
mHRTF	0.00	0.05	0.28	-	0.30	0.10	0.00
N1N2P1	0.00	0.13	0.45	-	0.50	0.20	0.00
N1N2P1+P2	0.00	0.05	0.43	-	0.50	0.15	0.00

D. 頭内定位率

被験者 YSD が 60° の実測 HRTF に対して 1 度だけ頭内に音像を知覚したことを除き, いずれの HRTF においても頭内定位は生じなかった.

3.2.2. 準 FEC 条件

A. 回答分布

FIG. 8 に実測 HRTF, N1N2P1, N1N2P1+P2 に対する 4 人の被験者の回答を示す.

被験者 MKI では, 実測 HRTF に対する回答はほぼ対角線上に分布した. N1N2P1 の回答は, 90° と 120° 以外では実測 HRTF と同様であった. 90° と 120° では回答の分散は実測 HRTF より大きい. N1N2P1+P2 の回答は 120° を除き実測 HRTF と同様であった. 120° では上方寄りに回答する傾向があった. 90° では, 回答の分散は N1N2P1 より小さく, 実測 HRTF と同様であった.

被験者 OIS では, 実測 HRTF に対する回答は逆 S 字カーブを描き, 30° , 60° ではやや上方に, 90° ではやや後方に回答する傾向があった. N1N2P1 の回答は, 0° , 180° では実測 HRTF と同様であった. しかし, 90° , 120° ではやや後方に回答し, 30° , 150° では回答の分散は大きい. N1N2P1+P2 の回答は, 90° , 150° を除き実測 HRTF と同様であった. 90° , 150° では N1N2P1 と同様の分布であった.

被験者 OTK では, 実測 HRTF に対しては 0° , 180° では目標方向に回答したが, 30° では前方から後方, 60° から 150° では上方から後方の広い範囲に回答が分布した. N1N2P1 の回答は, 0° , 180° では実測 HRTF と同

様であった。しかし 30°では回答の分散は実測 HRTF と比べて大きい。60°から 150°では後方に回答した。N1N2P1+P2 の回答は 30°を除き実測と同様であった。30°では回答の分散は実測 HRTF よりも大きい。

被験者 YSD では、実測 HRTF に対する回答は逆 S 字カーブを描き、30°、60°ではやや上方に、90°ではやや後方に回答する傾向があった。N1N2P1 の回答は 60°、90°を除き実測 HRTF と同様であった。60°、90°では回答は上方から後方に広く分布した。N1N2P1+P2 の回答はすべての目標方向で N1N2P1 と同様であった。ただし、90°の回答の分散はやや狭まり、150°の回答の分散はやや広がった。

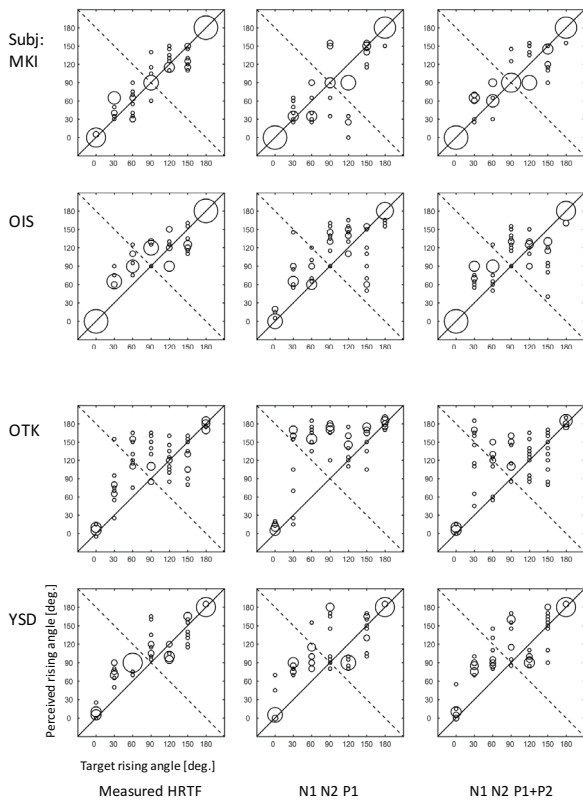


FIG. 8. Responses to the measured HRTF, N1N2P1, and N1N2P1+P2 under the quasi-FEC condition.

B. 平均上昇角定位誤差

各 HRTF と各目標方向の平均上昇角定位誤差(E_v)および、パラメトリック HRTF と実測 HRTF の E_v の差(ΔE_v)をそれぞれ求めた(FIGs. 9, 10).

実測 HRTF の E_v は、FEC 条件と同様に水平面上の 0°, 180°では小さく(4.7°, 2.1°), 水平面から上昇した 30°から 150°では大きい(16.2°-37.6°)傾向がある。N1N2P1, N1N2P1+P2 でも同様の角度依存性がみられた。しかし、これらの E_v は実測 HRTF より大きい。N1N2P1+P2 の E_v は 30°を除くすべての目標方向で N1N2P1 より小さい。

特に 60°, 90°, 120°では、N1N2P1 と比較してそれぞれ 13.4°, 17.5°, 11.8°減少した。

N1N2P1 と実測 HRTF の E_v の差(ΔE_r)は 0°, 150°, 180°では小さく(3.2°, 2.7°, 2.2°), 他の上昇角においては 9.9°から 23.5°であった。N1N2P1+P2 の ΔE_v は 30°では 19.6°であったが、他の方向では 10°未満であった。

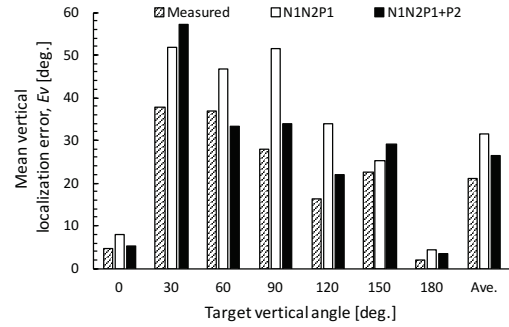


FIG. 9. Mean vertical localization error, E_v , for each HRTF and target vertical angle under the quasi-FEC condition.

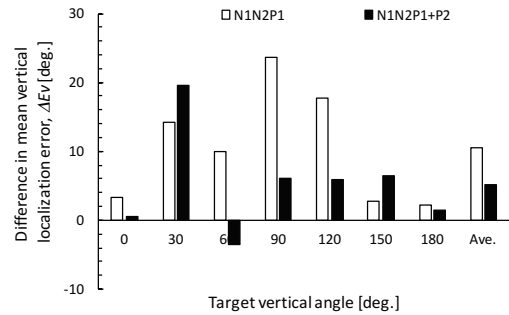


FIG. 10. Difference in E_v between the parametric HRTF and the measured HRTF, ΔE_v , under the quasi-FEC condition.

C. 前後誤判定率

TABLE II に前後誤判定率を示す。0°と 180°では、いずれの HRTF においても前後誤判定率は 0%であった。しかし、他の 4 方向においては、N1N2P1 と N1N2P1+P2 の前後誤判定率は実測 HRTF のそれより高い。

TABLE III にカイ 2 乗検定の結果を示す。120°において実測 HRTF と N1N2P1 の間に、30°において実測 HRTF と N1N2P1+P2 の間に統計的有意な差($p < 0.05$)が認められた。

TABLE II Ratio of front-back confusion for each subject, HRTF, and target vertical angle under the quasi-FEC condition.

HRTF	Target vertical angle (°)						
	0	30	60	90	120	150	180
mHRTF	0.00	0.08	0.40	-	0.05	0.05	0.00
N1N2P1	0.00	0.20	0.53	-	0.25	0.10	0.00
N1N2P1+P2	0.00	0.28	0.45	-	0.18	0.13	0.00

TABLE III Results of chi-square tests for the ratio of front-back confusion under the FEC condition. *: $p < 0.05$, **: $p < 0.01$.

Comparison between	Target vertical angle (°)						
	0	30	60	90	120	150	180
measured and N1N2P1					*		
measured and N1N2P1+P2		*		-			
N1N2P1 and N1N2P1+P2				-			

D. 頭内定位率

3人の被験者 (MKI, OIS, OTK)では、いずれの HRTF においても頭内定位は生じなかった。しかし被験者 YSD では、N1N2P1 (30°, 90°で1回, 60°で3回)および N1N2P1+P2 (30°, 60°, 90°, 120°で1回)で頭内定位が生じた。

4. 音像定位実験 2

音像定位実験 1 では、N1N2P1 に P2 を加えることにより正中面上方の音像定位精度が向上することを示した。音像定位実験 2 の目的は、P2 が正中面上方の音像定位で担う役割を解明することである。

4.1. 実験方法

実験方法は音像定位実験 1 と同様である。ただし、被験者本人の 90°の(1) P1, (2) P2, (3) P1P2 で再構成した3種類のパラメトリック HRTF を用いた。

FIG. 11 に実測 HRTF と P1, P2, P1P2 で再構成したパラメトリック HRTF の例を示す。パラメトリック HRTF は再構成したピークを精度よく再現し、その他の帯域のスペクトルは平坦である。

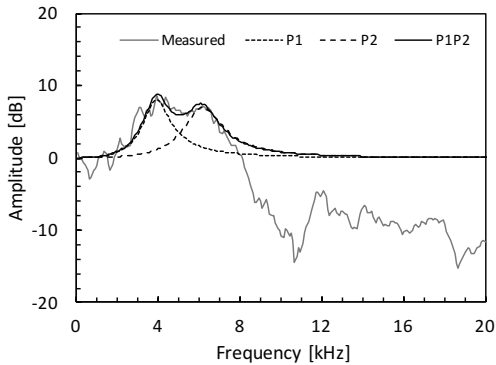


FIG. 11. Examples of a measured HRTF and parametric HRTFs (P1, P2, and P1P2) for a vertical angle of 90°.

4.2. 実験結果

4.2.1. 回答分布

FEC 条件での P1, P2, P1P2 に対する回答分布を FIG. 12 に示す。音像定位実験 1 で求めた目標方向 90° における実測 HRTF, N1N2P1, N1N2P1+P2 の回答分布も併せて示す。N1N2P1+P2 の回答分布は、3.2.1.項で示したように MKI でやや分散が大きくなるものの、実測 HRTF のそれと同様である。

一方、P1, P2, P1P2 に対する回答は被験者内ではほぼ同様に分布した。すなわち、MKI と OIS は前方から後方、OTK は上方から後方、YSD は斜め後ろから後方に広く分布した。いずれの被験者においても、上方に回答が集中することはなかった。

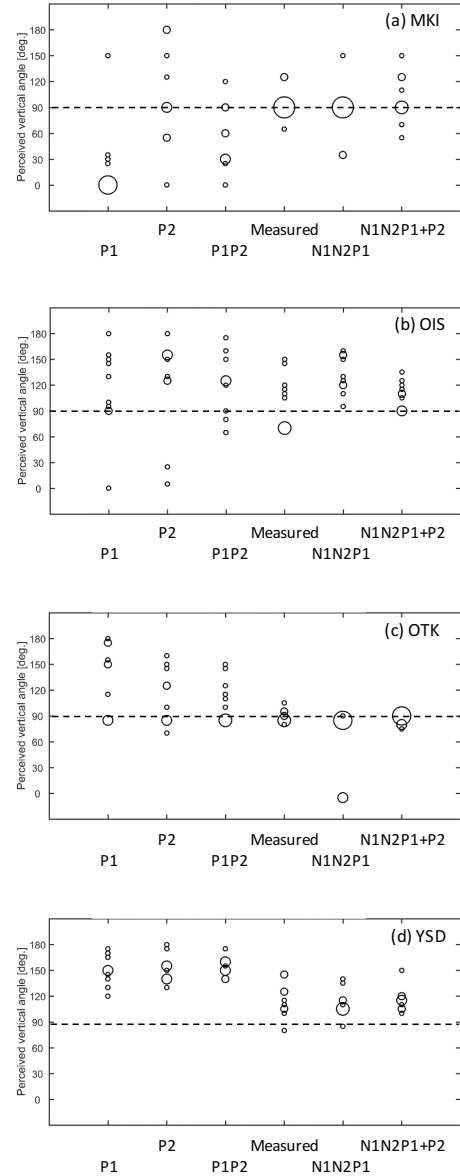


FIG. 12. Responses to P1, P2, and P1P2 under the FEC condition. For comparison, the responses to the measured HRTF, N1N2P1, and N1N2P1+P2 for a target vertical angle of 90° obtained in localization test 1 are also shown.

準 FEC 条件での回答分布を FIG. 13 に示す。各被験者の回答分布を Fig.13 に示す。N1N2P1+P2 では、3.2.2.項で示したように、OIS でやや分散が大きくなるものの、実測 HRTF と同様の回答分布であった。

一方、P1, P2, P1P2 に対する回答は、FEC 条件と同様に、被験者内ではほぼ同様に分布した。MKI は前方か

ら後方, OIS は前方から斜め後ろに広く分布した. OTK と YSD は斜め後ろから後方に分布した. いずれの被験者においても, 上方に回答が集中することはなかった.

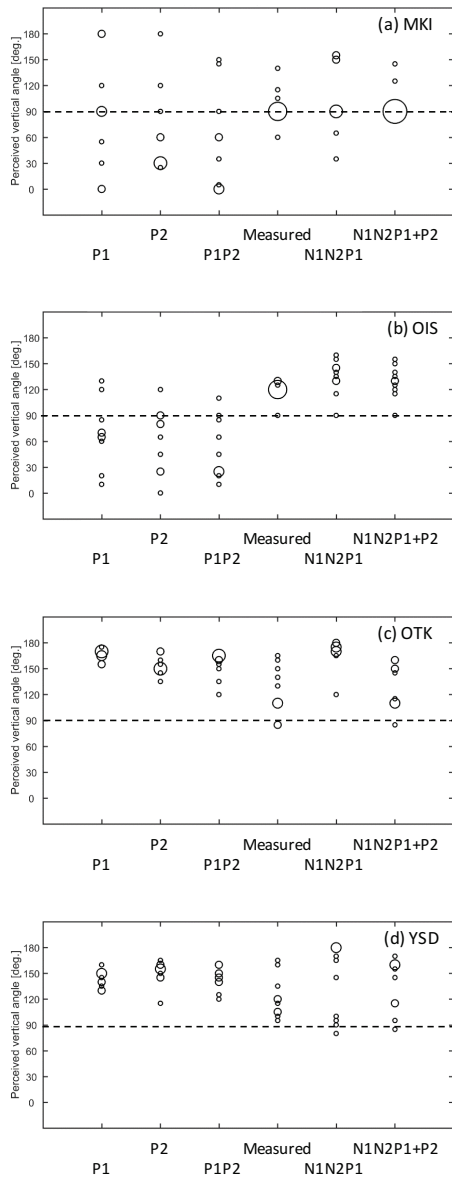


FIG. 13. Responses to P1, P2, and P1P2 under the quasi-FEC condition.

4.2.2. 平均上昇角定位誤差

平均上昇角定位誤差(E_v)を Fig.14 に示す.

FEC 条件(a)においては, N1N2P1+P2 の $E_v(17.7^\circ)$ は実測 HRTF(18.1°)と同等であった. しかし, P1, P2, P1P2 の $E_v(56.7^\circ, 48.8^\circ, 42.1^\circ)$ は実測 HRTF と比較して顕著に大きい値であった.

準 FEC 条件(b)においては, 実測 HRTF と N1N2P1+P2 の E_v はそれぞれ $27.9^\circ, 34.4^\circ$ であった. P1, P2, P1+P2 の

$E_v(52.9^\circ, 51.6^\circ, 54.6^\circ)$ は実測 HRTF と比較して顕著に大きい値であった.

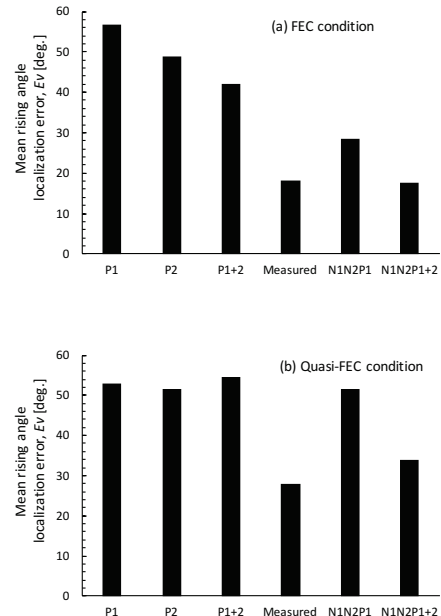


FIG. 14. Mean vertical localization error, E_v , under (a) the FEC condition and (b) the quasi-FEC condition at a target vertical angle of 90° .

4.2.3. 頭内定位率

P1, P2, P1P2 のいずれにおいても, どの被験者も頭内定位は生じなかった.

5. 考察

5.1. 従来の研究との比較

先に述べたように, Hebrank and Wright (1974)はフィルタ処理なしと比較して, ある方向に 1.5 倍以上の回答が得られたフィルタをスペクトラルキューと定義した. 彼らは, 上方のキューは 7 kHz から 9 kHz の間の 1/4 オクターブのピークであると主張した. このピークは P2 と一致し, さらに Blauert (1969/70)が提案した上方の方向決定帯域とも一致する.

しかし, 第 3,4 章の結果は, 広帯域信号に対しては P2 を再現しただけでは上方に音像定位しないことを明確に示している. つまり, Hebrank and Wright の主張するピークは上方のスペクトラルキューではない.

5.2. 正中面上方における P2 の役割

FIG. 15 は, FEC 条件の目標方向 90° において, N1N2P1 で上方と前方に回答し, N1N2P1+P2 では上方に回答した被験者 OTK の N1, N2, P1, P2 の周波数とレベルの関係を示したものである. 白丸は 0° , 黒丸は 90° の N1, N2, P1, P2 を示す. 2本の破線は, Moore *et al.*

(1989)が中心周波数 8 kHz, 帯域幅が中心周波数の 25% のノッチを用いて求めた 3 人の被験者のノッチ検知閾の最大値と最小値を示している. すなわち, ノッチレベルが -9dB 以上ではすべての被験者がノッチを検知できず, -20dB 以下ではすべての被験者が検知した.

0° の $N1, N2$ のレベルはともに検知閾を超えている. しかし, 90° の $N1$ は検知できないレベルであり, $N2$ は被験者により検知の可否が分かれるレベルである.

周波数軸上でみると, 90° ではノッチとピークは周波数の低い方から $P1, P2, N1, N2$ の順で並んでいる. $P2$ を再現しない場合, $P1$ は $N1$ から離れた周波数にあるため対比効果は期待できない. しかし $P2$ を再現すると, $P2$ からみた $N1$ の相対的なレベルは -14.7dB となる. これは被験者によってはノッチの検知が可能なレベルである.

以上より, $P2$ それ自体はスペクトラルキューではないが, $N1$ を強調することにより, 正中面上方の音像定位精度を向上させる重要な役割を果たしている.

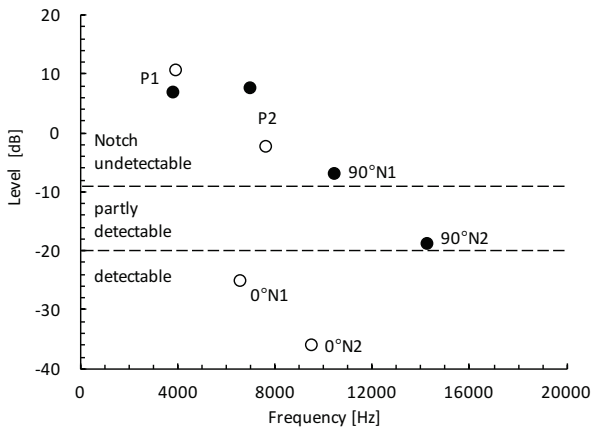


FIG. 15. Relationship between the frequencies and the levels of $N1, N2, P1,$ and $P2$ for subject OTK. Open and closed circles denote $N1, N2, P1,$ and $P2$ for 0° and 90° , respectively. Broken lines indicate the maximum and minimum thresholds of three subjects for the detection of the notch, the center frequency of which is 8 kHz and the bandwidth of which is 25 % of the center frequency, as reported by Moore *et al.* (1989).

5.3. FEC 条件と準 FEC 条件の比較

FEC 条件と準 FEC 条件の平均上昇角定位誤差 (E_v) を比較する. FIG. 16 に各 HRTF と各条件の E_v を示す.

実測 HRTF(a) では, $0^\circ, 180^\circ$ においては, FEC 条件と準 FEC の E_v の差は小さい ($1.5^\circ, 0.3^\circ$). したがって, 前後方向の音像定位においては, $M \times PTF$ の補正は必要ではない. しかし, $30^\circ, 60^\circ, 90^\circ$ においては, 準 FEC 条件の E_v は FEC 条件と比較して顕著に大きく, その差はそれぞれ $16.4^\circ, 13.0^\circ, 9.8^\circ$ であった. 一方, $120^\circ, 150^\circ$

では FEC 条件よりも準 FEC 条件の方が E_v はやや小さい. これは, 準 FEC 条件では FEC 条件と比較して後方に知覚する傾向があることに起因する (FIGs. 5, 8 参照). 後方に音像を知覚する本質的な理由は不明であるが, Fig. 4 に示した $M \times PTF$ の peak-to-peak レンジの広さが根底にある理由の一つかもしれない. 準 FEC 条件において上方から後方への回答が頻繁にみられた被験者 OTK の $M \times PTF$ の peak-to-peak レンジは他の被験者と比べて大きく, 定位誤差の大きさと対応している.

$N1N2P1(b)$ でも $0^\circ, 180^\circ$ では FEC 条件と準 FEC の E_v の差は小さい ($-1.5^\circ, 2.1^\circ$). しかし, $30^\circ, 60^\circ, 90^\circ$ では準 FEC 条件の E_v は FEC 条件と比較して顕著に大きく, その差はそれぞれ $16.5^\circ, 14.9^\circ, 23.0^\circ$ であった.

$N1N2P1+P2(c)$ では, $30^\circ, 90^\circ$ を除き FEC 条件と準 FEC の E_v の差は小さい. しかし, $30^\circ, 90^\circ$ では準 FEC 条件の E_v は FEC 条件と比較して顕著に大きく, その差はそれぞれ $27.9^\circ, 16.3^\circ$ であった.

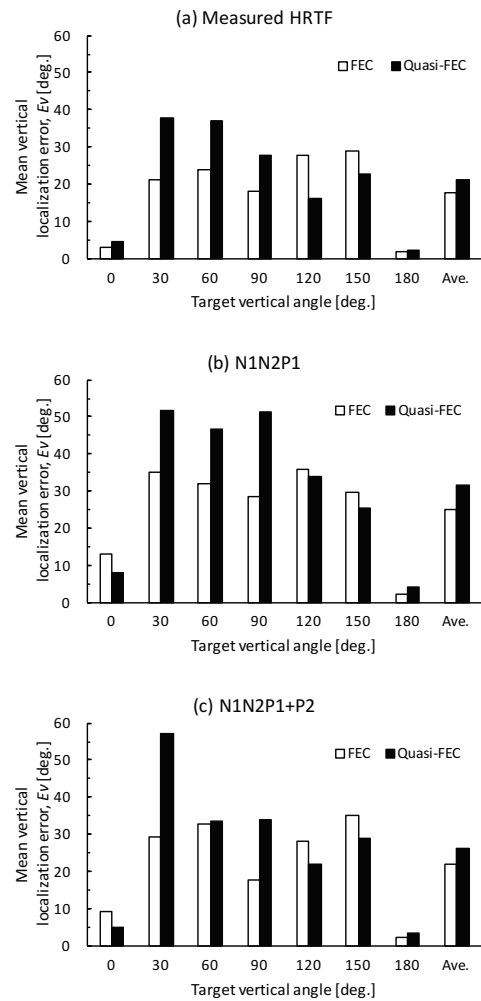


FIG. 16. Mean vertical localization error, E_v , for the FEC and quasi-FEC conditions for (a) the measured HRTF, (b) $N1N2P1,$ and (c) $N1N2P1+P2.$

以上をまとめると、実測 HRTF, N1N2P1, N1N2P1+P2 のいずれにおいても、 0° および 120° から 180° では準 FEC 条件は FEC 条件と同等の音像定位性能を有している。しかし 30° から 90° においては、準 FEC 条件の E_v は FEC 条件と比較して顕著に大きく、 $M \times PTF$ の補正が必要である。

6. 結論

本論文では、音像定位実験 1 により、N1, N2, P1 に P2 を加えて再構成したパラメトリック HRTF (N1N2P1+P2)により正中面上方で実測 HRTF と同等の音像定位精度が得られるか否かを検証した。さらに、音像定位実験 2 により、正中面上方の音像定位における P2 の役割を考察した。

FEC 条件における結果を以下に示す。

(1) N1N2P1 に P2 を加えることにより、正中面の 90° および 120° で音像定位性能が向上した。

(2) N1N2P1+P2 と実測 HRTF の間の平均上昇角定位誤差の差は上半球正中面の全 7 方向において 10° 未満であった。

(3) P1, P2, P1P2 を再現しても音像は上方に知覚されなかった。言い換えれば、これらのスペクトラルピークは上方のスペクトラルキューではない。

(4) 0° の N1 と N2 のレベルはともにノッチの検知閾を超えるが、 90° の N1 は検知できない。しかし、P2 を加えることにより、P2 からみた N1 の相対レベルは検知可能となる。

(5) 以上より、P2 は N1 を強調することにより、正中面上方の音像定位精度を向上する重要な役割を果たしていると考えられる。

さらに、準 FEC 条件における結果を以下に示す。

(6) N1N2P1 に P2 を加えることにより、正中面の 60° から 120° で音像定位性能が向上した。

(7) 0° および 120° から 180° では、FEC 条件と同等の音像定位性能が得られた。しかし、 30° から 90° では音像定位性能は FEC 条件と比較して低下した。

謝 辞

本研究の一部は科研費(基盤研究(A) 15H01790)により実施した。また、神戸大学の森本政之名誉教授には有益な議論をしていただいた。記して感謝いたします。

文 献

Asano, F., Suzuki, Y., and Sone, T. (1990). "Role of spectral cues in median plane localization," J Acoust. Soc. Am. **88**, 159–168.

Blauert, J. (1969/70). "Sound localization in the median plane," Acustica, **22**, 205–213.

Butler, A., and Belendiuk, K. (1977). "Spectral cues utilized in the localization of sound in the median sagittal plane," J. Acoust. Soc. Am. **61**, 1264–1269.

Carlile, S., Leong, P., and Hyams, S. (1997). "The nature and distribution of errors in sound localization by human listeners," Hear. Res. **114**, 179–196.

Gardner, B., and Gardner, S. (1973). "Problem of localization in the median plane: effect of pinna cavity occlusion," J. Acoust. Soc. Am. **53**, 400–408.

Hebrank, J., and Wright, D. (1974). "Spectral cues used in the localization of sound sources on the median plane," J. Acoust. Soc. Am. **56**, 1829–1834.

Iida, K., Yairi, M., and Morimoto, M. (1998). "Role of pinna cavities in median plane localization," Proceedings of 16th International Congress on Acoustics, 845–846.

Iida, K., Itoh, M., Itagaki, A., and Morimoto, M. (2007). "Median plane localization using parametric model of the head-related transfer function based on spectral cues," Appl. Acoust. **68**, 835–850.

Iida, K., and Ishii, Y. (2011). "Individualization of the head-related transfer functions in the basis of the spectral cues for sound localization," in Principles and Applications of Spatial Hearing, edited by Y. Suzuki, D. Brungard, Y. Iwaya, K. Iida, D. Cabrera, and H. Kato (World Scientific, Singapore), pp. 159–178.

Iida, K., Nishioka, S., and Ishii, Y. (2014). "Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae," J. Acoust. Soc. Am. **136**, 317–333.

Kulkarni, A., and Colburn, H. S. (1998). "Role of spectral detail in sound-source localization," Nature **396**, 747–749.

Lopez-Poveda, E. A., and Meddis, R. (1996). "A physical model of sound diffraction and reflections in the human concha," J. Acoust. Soc. Am. **100**, 3248–3259.

Majdak, P., Goupell, M. J., and Laback, B. (2010). "3-D localization of virtual sound sources: Effects of visual environment, pointing method, and training," Atten. Percept. Psychophys. **72**, 454–469.

Møller, H., Hammershøi, D., Jensen, C. J., and Sørensen, M. F. (1995). "Transfer characteristics of headphones measured on

human ears,” *J. Audio Eng. Soc.* **43**, 203–217.

Moore, B. C. J., Oldfield, R., and Dooley, G. J. (1989). “Detection and discrimination of peaks and notches at 1 and 8 kHz,” *J. Acoust. Soc. Am.* **85**, 820–836.

Morimoto, M., and Aokata, H. (1984). “Localization cues of sound sources in the upper hemisphere,” *J. Acoust. Soc. Jpn. (E)* **5**, 165–173.

Musicant, A., and Butler, R. (1984). “The influence of pinnae-based spectral cues on sound localization,” *J. Acoust. Soc. Am.* **75**, 1195–1200.

Raykar, V. C., Duraiswami, R., and Yegnanarayana, B. (2005). “Extracting the frequencies of the pinna spectral notches in measured head related impulse responses,” *J. Acoust. Soc. Am.* **118**, 364–374.

Shaw, E. A. G., and Teranishi, R. (1968). “Sound pressure generated in an external-ear replica and real human ears by a nearby point source,” *J. Acoust. Soc. Am.* **44**, 240–249.

Takemoto, H., Mokhtari, P., Kato, H., Nishimura, R., and Iida, K. (2012). “Mechanism for generating peaks and notches of head-related transfer functions in the median plane,” *J. Acoust. Soc. Am.* **132**, 3832–3841.