

受聴者の耳介形状による頭部伝達関数の個人化：再考

飯田 一博[†] 石井 要次[‡] 西岡 伸介[‡]

[†] 千葉工業大学工学部, [‡] 大学院工学研究科 〒275-0016 千葉県習志野市津田沼 2-17-1

E-mail: [†] kazuhiko.iida@it-chiba.ac.jp, [‡] {s1179501rq, s1272022yk}@s.chibakoudai.jp

あらまし 本研究では、それぞれの受聴者に適した頭部伝達関数(HRTF)を推定する方法を提案する。HRTFにおける最も周波数の低いスペクトラルノッチ(N1・N2)は上昇角知覚において重要な役割を果たすが、提案法は受聴者の耳介形状からその N1・N2 周波数を推定する。さらに、推定値に最も近い周波数の N1・N2 をもつ HRTF を best-matching HRTF としてデータベースから選出する。提案法の有効性を確認するために4人の被験者を用いて上半球正中面音像定位実験を行った。その結果、best-matching HRTF の音像定位精度は、正面および後方のターゲット方向に対しては、全ての被験者において被験者本人の HRTF と同等であった。一方、上方のターゲット方向に対しては、一部の被験者において被験者本人の HRTF と比較して精度が低下した。

キーワード 頭部伝達関数, 個人化, 個人適応, 耳介, スペクトラルノッチ

Personalization of head-related transfer functions based on the anthropometry of listener's pinna: revisited

Kazuhiro IIDA[†] Yohji ISHII[‡] and Shinsuke NISHIOKA[‡]

{[†] Faculty of 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, s1272022yk}@s.chibakoudai.jp

Abstract The present study proposes a method for estimating the appropriate HRTFs for an individual listener. The proposed method estimates the frequencies of the two lowest spectral notches (N1 and N2), which play an important role in vertical localization, in the HRTF of an individual listener by anthropometry of the listener's pinnae. The best-matching HRTFs, of which N1 and N2 are the closest to the estimates, are then selected from an HRTF database. In order to examine the validity of the proposed method, localization tests in the upper median plane were performed using four subjects. The results revealed that the best-matching HRTFs provided approximately the same performance as the listener's own HRTFs for the target directions of front and rear for all the four subjects. For the upper target directions, however, the performance of the localization for some of the subjects degraded.

Keyword head-related transfer function, personalization, individualization, spectral notch, best-matching HRTF

1. はじめに

受聴者本人の頭部伝達関数(HRTF: head-related transfer function)を再現すれば精度の高い音像制御が実現できる(Morimoto and Ando, 1980)。しかし、HRTFの測定には特別な装置や多くの時間を要するため、任意の受聴者の任意の方向のHRTFを測定することは事実上不可能である。

Soitek and Genuit (1999) は、HRTFの個人化のゴールの1つとして”a person who enters a multimedia shop is scanned by a camera and some instants later his/her individual HRTF set is ready for use in advanced 3D applications”というBlauertのシナリオを示した。しかし、このシナリオは未だに実現されていない。

これまで、個人化HRTFを得る方法を確立するため多くの取組みがなされてきた。そのような方法の1つとして、HRTFをいくつかの主成分に分解するPCA(principal component analysis)がある(Kistler and Wightman, 1992; Middlebrooks and Green, 1992)。また、各主成分の係数を受聴者の耳介形状から求める試みもある(Rodriguez and Ramirez, 2005; Hu *et al.*, 2008; Xu *et al.*, 2008; Hugeng *et al.*, 2010)。しかし、これらの考え方に基づいてHRTFを個人化し、その性能を音像定位実験で検証したという報告はない。

一方、数値計算によるHRTFの推定も盛んに取り組まれている。従来、BEM(boundary element method)が多くの研究で使われてきた(Katz, 2001; Kahana and

Nelson, 2006; Kreuzer *et al.*, 2009). しかし、最近では BEM よりも演算速度の早い FDTD (finite-difference time-domain)法が主流になりつつある。この方法により、HRTF の基本的なスペクトル特性は耳介だけで生成されていることが示されている(Takemoto *et al.*, 2012)。しかし、現時点では BEM にせよ FDTD 法にせよ任意の受聴者に適用することはできない。なぜなら、受聴者の耳介形状をデジタル化するには fMRI(functional magnetic resonance imaging system)などのような特別な装置が必要となるからである。

受聴者自身が試聴することによって適した HRTF を選択する方法についても多くの研究がみられる。Middlebrooks (1999a,b)は HRTF の方向依存成分である DTF (directional transfer functions)の個人差は周波数軸状での scaling で削減できると報告した。さらに、Middlebrooks (2000) は被験者本人の DTF および適切に frequency-scaled した DTF の quadrant errors はそれぞれ 15.6%および 14.7%であることを示した。ここで、quadrant error は上下もしくは前後方向における 90°以上の誤差と定義される。しかし、受聴者に適した scale factor を見いだすには、1 ブロックで 20 分要する試聴実験が 1 から 3 ブロック必要となる。Iwaya (2006)はトーナメント方式の試聴実験を提案しているが、32 の HRTF セットから受聴者に適したセットを選出するのに 15 分要すると報告している。この場合、データベースに含まれる HRTF が増えれば増えるほど試聴実験に要する時間も増えることに留意しなければならない。

Zotkin *et al.* (2003) は受聴者の耳介形状を計測し、データベースの中から最も近い形状の耳介の HRTF を選出する方法を提案している。彼らは提案法で選出した HRTF と KEMAR 疑似頭の HRTF を用いて、前半球面において音像定位実験を行った。その結果、提案法による音像定位精度の改善は、8 人の被験者の平均値で僅か 1.9°であった。全ての耳介形状の HRTF への寄与は等しいとは考えられないため(Takemoto *et al.*, 2012), 彼らの方法の効果が少ないことは不思議ではない。各耳介形状パラメータの音像定位の手掛かりへの寄与を考慮することにより、高い音像定位精度をもたらす HRTF をデータベースから選出することができるようになると著者らは考える。

飯田ら(2013)は、音響測定や試聴実験をすることなしに受聴者に適した HRTF を提供する方法を提案した。この方法は、受聴者の耳介形状からその受聴者の上昇角知覚のスペクトラルキューを推定し、さらに推定値に最も近いスペクトラルキューをもつ HRTF をデータベースから選出するというものである。

その後、われわれはこの方法にいくつかの改良を加えた。その結果、より少ない耳介形状パラメータから

受聴者に適した HRTF を提供することが可能になったので、ここに報告する。また、本研究で用いた HRTF データベースを公開したので、その情報を併せて示す。

2. スペクトラルキューの個人差

2.1. 推定すべきスペクトラルキュー

5kHz 以上のスペクトラルピークとノッチは仰角知覚において重要な役割を果たしている (Hebrank and Wright, 1974; Butler and Belendiuk, 1977; Mehrgardt and Mellert, 1977; Musicant and Butler, 1984)。

Kulkarni and Colburn (1998) は被験者の水平面内 (0, ±45, 180°)の HRTF を 7 段階で平滑化して音像定位実験を行い、スペクトルの細部は重要ではないことを示した。また、過度の平滑化を行うと音像が水平面から上昇することを示した。

Iida *et al.* (2007)は、音源の仰角に依存せず出現する (Shaw and Teranishi, 1968) 4kHz 付近のピーク(P1)を下限として、受聴者の実測 HRTF からスペクトラルピークとノッチを抽出した。彼らは上半球正中面で音像定位実験を行い、P1 とそれに続く 2 つのノッチ(N1, N2)のみで再構成した簡略化 HRTF でも実測 HRTF と同等の音像定位精度が得られることを示した。さらに、彼らは N1・N2 周波数は音源の上昇角に強く依存することを示した。Fig.1 はある被験者の正中面内の HRTF の振幅スペクトルである。図中、破線は 4 次多項式による N1・N2 近似曲線、実線は P1 の近似直線である。

前述の Kulkarni and Colburn (1998)の過度の平滑化条件では音像が上昇して知覚されたが、この場合のスペクトルをみると N2 が消失しており、N2 の重要性を示唆している。

以上より、N1・N2 は上昇角知覚における手掛かりとして重要な役割を果たしていると考えられる。また、P1 は聴覚システムが耳入力信号から N1・N2 を分析するためのリファレンス情報として機能していると解釈できる。

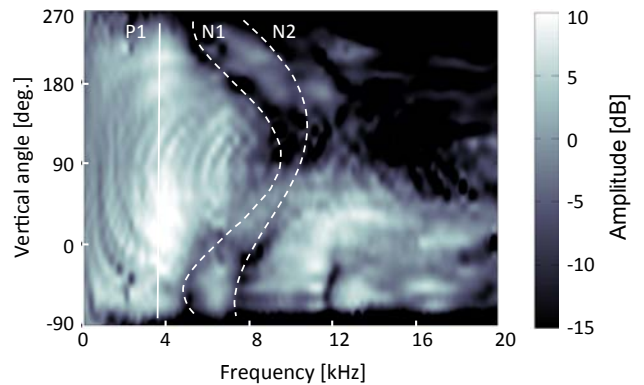


Figure 1 Amplitude spectrum of HRTF in the median plane.

2.2. N1・N2・P1 周波数の個人差

このように上昇角知覚において重要な情報は HRTF の N1・N2・P1 周波数に絞ることができる。しかし、これらの周波数には個人差があるため、他人の HRTF を用いると頻りに前後誤判定、上昇角誤差、頭内定位などが生じる。そこで、正面方向の HRTF における N1・N2・P1 周波数の個人差を測定により求めた。

2.2.1. HRTF の測定

無響室において、28 人の日本人成人の上半球正中面 7 方向(30°間隔)の HRTF を測定した。音源用スピーカ (FOSTEX FE83E) から swept sine wave(2¹⁸ サンプル) を放射し、被験者の外耳道入口に装着した耳栓型マイクロホン (Iida *et al.*, 2007) で受音した (blocked-entrance condition)。サンプリング周波数は 48kHz、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。HRTF は以下の式により求めた。

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

ここで、 $F(\omega)$ は自由空間において被験者のいない状態で音源から被験者の頭部中心位置に相当する位置までのインパルス応答 $f(t)$ のフーリエ変換であり、 $G_{l,r}(\omega)$ は被験者の外耳道入口で測定したインパルス応答 $g(t)$ のフーリエ変換である。

2.2.2. N1・N2・P1 の抽出

HRTF の測定を行った 28 人の 56 耳について、正面方向の HRTF の N1・N2・P1 を抽出した。N1・N2・P1 は耳介で発生しているため (Shaw and Teranishi, 1968; Lopez-Poveda and Meddis, 1996; Takemoto *et al.*, 2012)、頭部インパルス応答 (HRIR: head-related impulse response) の初期部分から以下のアルゴリズムにより抽出した。

- 1) HRIR の振幅の絶対値が最大となるサンプルを検出する。
 - 2) HRIR を 4 次 96 ポイントの Blackman-Harris 窓で切り取る。ただし、1) で検出した最大サンプルを時間窓の中心に合わせる。
 - 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) で切り取った 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 により 512 ポイントの配列の振幅スペクトルを求める。
 - 5) 差分法により極大値および極小値を検出する。最も周波数の低い極大値を P1 とし、それ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。
- 結果として、54 耳(26 人の両耳と 2 人の片耳)の N1・N2・P1 を抽出した。残りの 2 耳については、スペクトラルノッチが 1 つしか抽出できなかったため、以降の

分析からは除外した。

54 耳の N1・N2・P1 周波数の分布範囲はそれぞれ、5,719–9,563Hz (0.74 oct.)、8,250–13,500Hz (0.71 oct.)、3,469–4,313Hz (0.31 oct.)であった。P1 周波数の個人差は N1 や N2 のそれよりも小さく、N1 と N2 の周波数の分布範囲は互いに重なっている。

上昇角知覚における正面方向の P1 周波数の丁度可知差異 (JND: just-noticeable difference) は 0.35 oct. (上側周波数) および 0.47 oct. (下側周波数) である (西岡他, 2013)。したがって、P1 周波数の個人差は上昇角知覚において影響しないとみなせる。そこで、以降 P1 は議論の対象から除外する。一方、N1・N2 の JND は 0.1–0.2 oct. 程度であり (Iida and Ishii, 2011b)、N1・N2 周波数の個人差 (0.74 および 0.71 oct.) は上昇角知覚において十分影響を及ぼすと考えられる。

3. 耳介形状の個人差

Takemoto *et al.* (2012) は 4 人の被験者の頭部形状から FDTD 法により HRTF を算出し、4 人の N1・N2 周波数はそれぞれ異なるが、ピーク・ノッチの基本パターンは耳介により生じることを示した。さらに、N1 周波数では耳介の腔に 1 つもしくは 2 つの腹と 1 つの節が出現していると報告した。これらの知見は、N1・N2 周波数の個人差は受聴者の耳介形状に起因することを示唆している。

そこで、本研究では Algazi *et al.* (2001) と同様の 10 種類の耳介形状パラメータを採り上げ (Fig.2)、その個人差を分析した。ただし、彼らが用いたパラメータの 1 つである q2 (pinna flare angle) は実際には計測が困難であるので、その代わりとして x_4 (width of the helix) を用いた。

54 耳の耳型より 9 種類の耳介形状パラメータ (x_1 – x_8 および x_d) をノギスで計測した。耳介の傾き (x_a) は被験者の横顔の写真から計測した。計測結果を Table 1 に示す。各パラメータの分布の幅は 10–25 mm であった。また、耳介の傾き (x_a) は 4–40° であった。

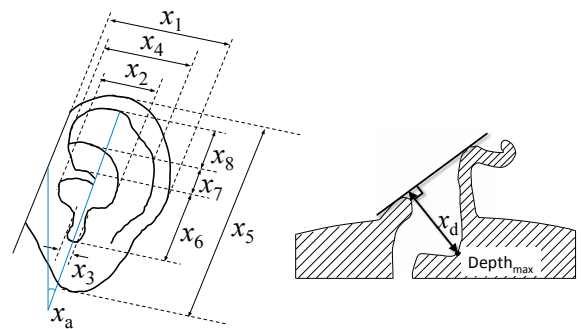


Figure 2 Ten anthropometric parameters of the pinna.

Table 1 Measured ten pinnae dimensions for 54 ears (mm).

Pinna	Width of				Length of				Depth of concha	Tilt of pinna (°)
	pinna	concha	incisura intertragica	helix	pinna	concha	cymba conchae	scapha		
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_d	x_a
1	37.0	20.1	8.3	29.7	69.1	21.0	9.3	16.8	15.2	18
2	35.7	19.6	6.5	29.4	72.1	21.5	9.8	15.5	14.9	24
3	33.0	19.7	5.9	27.0	66.3	21.8	6.4	19.8	15.8	29
4	37.5	15.8	5.5	24.7	75.8	19.3	9.9	18.8	13.8	30
5	35.9	18.6	9.0	24.7	73.8	22.9	6.2	19.6	11.9	16
6	34.2	19.5	9.8	26.3	72.0	22.3	9.3	17.9	12.1	17
7	34.7	14.9	5.3	20.1	71.3	21.4	5.9	20.8	16.7	7
8	31.9	16.8	6.8	25.2	70.1	20.2	4.5	21.4	17.2	12
9	34.4	17.8	7.0	25.7	71.4	22.3	5.3	19.4	13.3	17
10	35.0	18.8	9.0	26.1	63.4	22.2	4.7	20.8	14.1	18
11	34.4	18.2	8.3	24.8	64.4	20.4	7.1	20.8	15.0	19
12	38.0	17.6	8.3	28.3	67.0	17.7	8.8	19.2	12.5	23
13	35.1	19.1	8.4	26.3	68.5	17.1	8.1	19.8	13.8	7
14	34.8	18.3	8.7	22.2	66.4	18.8	6.6	22.5	14.1	23
15	35.2	18.7	8.3	27.1	67.0	18.6	8.3	22.0	13.8	24
16	36.1	16.5	5.5	27.3	64.3	19.3	5.5	17.3	14.0	30
17	37.5	16.6	5.5	29.7	66.4	18.7	8.0	17.2	13.8	40
18	36.6	21.2	8.4	27.0	67.2	19.3	7.5	18.4	11.3	27
19	36.1	19.8	7.5	28.1	66.4	19.2	8.5	18.3	11.1	21
20	35.7	15.6	5.8	25.7	69.2	20.5	8.6	20.3	13.7	28
21	35.8	15.4	6.4	26.8	70.8	20.9	9.4	21.2	13.1	36
22	36.1	20.5	6.7	27.0	64.0	21.0	6.0	20.2	14.8	32
23	33.4	16.5	5.9	27.5	63.8	22.5	4.3	21.1	13.8	27
24	43.8	20.2	8.3	31.6	78.3	21.9	8.0	24.1	14.1	27
25	42.1	19.2	7.3	31.8	77.8	19.0	9.3	23.1	14.4	20
26	36.9	14.8	6.7	27.5	72.8	24.1	7.5	19.4	13.0	23
27	36.1	17.4	7.5	24.3	70.4	22.9	9.3	17.4	13.8	31
28	31.7	16.6	8.9	19.0	69.2	23.3	7.7	15.9	16.0	25
29	32.8	16.7	8.7	18.1	69.3	20.4	9.1	17.1	15.7	24
30	36.0	20.1	10.8	28.3	64.0	22.0	5.8	16.7	13.8	27
31	34.0	21.8	11.9	26.9	64.4	22.0	10.3	14.2	13.1	36
32	40.6	19.1	10.2	28.0	75.8	25.1	2.6	21.1	14.1	23
33	36.4	18.4	7.6	26.2	83.2	24.9	4.7	22.1	13.8	29
34	31.2	19.3	10.9	21.6	63.5	20.4	5.9	15.4	13.3	13
35	32.5	19.9	10.1	23.0	65.1	21.3	8.2	15.9	11.7	14
36	33.1	21.0	9.0	23.1	58.2	18.4	3.5	14.0	9.9	32
37	32.7	20.4	7.2	26.6	59.1	17.8	5.1	16.9	9.7	16
38	35.7	19.7	7.4	29.0	66.9	20.2	5.3	17.6	11.8	19
39	35.0	21.8	7.3	28.8	68.9	19.5	7.7	20.0	12.5	28
40	34.5	17.5	8.4	24.3	66.1	23.0	5.9	16.5	14.9	29
41	34.5	19.9	8.1	25.3	68.2	23.1	6.3	17.9	15.0	17
42	35.2	19.0	10.0	26.6	69.9	21.6	4.0	20.3	12.6	18
43	34.5	19.1	6.1	28.3	72.5	21.0	6.0	20.7	15.2	20
44	37.4	20.9	7.5	29.4	73.8	19.4	9.7	23.2	13.6	33
45	35.3	20.0	11.2	25.2	63.4	24.1	3.5	18.5	15.6	14
46	33.2	18.5	9.9	22.8	68.7	24.5	3.9	16.5	17.6	10
47	32.9	18.5	10.4	24.0	62.4	21.2	5.8	18.9	14.8	4
48	35.9	19.2	9.0	26.1	63.2	20.1	7.1	17.1	14.0	17
49	32.9	20.0	9.3	23.2	64.5	23.5	2.9	13.2	17.6	14
50	37.3	18.1	8.3	23.9	65.7	21.8	4.7	14.9	17.4	15
51	35.7	18.1	8.8	26.2	66.7	22.3	6.1	18.8	13.4	20
52	38.4	17.9	10.1	26.4	65.4	20.4	5.6	20.8	11.3	9
53	36.9	19.9	11.5	28.8	66.8	21.0	6.5	19.4	13.6	39
54	38.0	21.7	8.1	28.2	68.8	22.8	7.5	21.4	14.8	28
Min	31.2	14.8	5.3	18.1	58.2	17.1	2.6	13.2	9.7	4
Max	43.8	21.8	11.9	31.8	83.2	25.1	10.3	24.1	17.6	40

4. 耳介形状による N1・N2 周波数の推定

4.1. 重回帰モデル

上記の 54 耳を用い、10 種類の耳介形状パラメータを説明変数、正面方向の N1・N2 周波数を目的変数として重回帰分析を行った(式(2)).

$$f(S)_{N1,N2} = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n + b \text{ [Hz]} \quad (2)$$

ここで、 S , a_i , b , and x_i はそれぞれ被験者、重回帰係数、定数、耳介形状パラメータを表す。

4.2. 重回帰モデルの精度

重回帰分析の結果、N1 および N2 の重相関係数は 0.84 および 0.87 であった。しかし、各説明変数の p 値については、N1 については x_1, x_4, x_7 で 0.05 を超えた。N2 については同様に x_1, x_2, x_3, x_4, x_a で 0.05 を超えた。

これらのパラメータは N1・N2 に寄与しているとは考えにくい。そこで、説明変数の個数を変えながら全ての説明変数の組み合わせについて重回帰分析を行い、用いた全ての説明変数の p 値が 0.05 未満かつ重相関係数が最大となる組み合わせを求めた。その結果、N1 については 6 種類 ($x_2, x_3, x_6, x_8, x_d, x_a$), N2 については 3 種類 (x_6, x_8, x_d) の耳介形状パラメータが選ばれた。つまり、N1 においては耳介の窪みの幅、長さ、深さと耳介の傾きが寄与し、N2 においては耳介の窪みの長さ、深さが寄与していると言える。重回帰係数、p 値、95% 信頼区間を Table 2 に示す。

Table 2 Multiple regression coefficients, p-values, and 95% confidence intervals of N1 and N2 for the front direction.

	Regression coefficient		P-value		95% confidence intervals			
	N1	N2	N1	N2	lower	upper	lower	upper
a_1								
a_2	116.9		1.6E-02		22.9	210.9		
a_3	-157.5		4.7E-03		-264.2	-50.8		
a_4								
a_5								
a_6	-183.4	-327.0	8.3E-05	2.9E-07	-269.1	-97.8	-438.0	-216.0
a_7								
a_8	-93.2	-245.0	2.3E-03	4.4E-08	-151.5	-34.9	-321.3	-168.6
a_d	-131.4	-172.8	4.0E-03	3.7E-03	-218.7	-44.2	-286.9	-58.7
a_a	-48.7		7.2E-07		-65.8	-31.6		
b	14906.4	23903.1	9.2E-14	2.0E-22	12019.9	17792.9	21079.9	26726.3

重回帰モデルで推定した N1・N2 周波数と実測 HRTF から抽出したそれらとの関係を Fig.3 に示す。また、重回帰モデルの統計量を Table 3 に示す。N1 および N2 の重相関係数はそれぞれ 0.81 および 0.82 であった。平均絶対残差はそれぞれ 0.07 および 0.08 oct. であった。絶対残差が JND 以内となる確率は N1・N2 共通に 91% であった。ここで、JND は先述の通り 0.1-0.2 oct. と考えられるので 0.15 oct. とした。

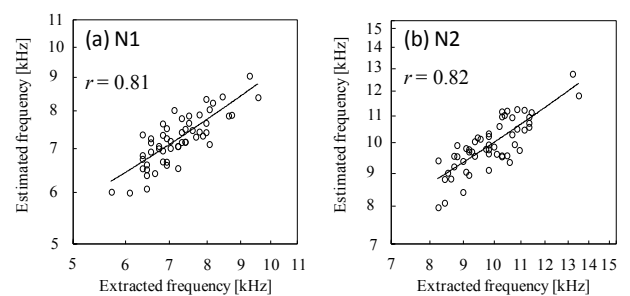


Figure 3 Relationship between the frequencies extracted from the measured HRIR and those estimated from the listener's anthropometric parameters for 54 ears. (a): N1, (b): N2.

Table 3 Statistics of the multiple regression models of N1 and N2 for the front direction.

	Correlation coefficient	Significance level	Absolute mean residual error [Hz]	Probability that residual error is less than 0.15 octaves [%]
N1	0.81	1.1E-09	357	91
N2	0.82	5.4E-12	550	91

さらに、多重共線性の有無を確認するために $x_2, x_3, x_6, x_8, x_d, x_a$ の間、および x_6, x_8, x_d の間で VIF (variance inflation factor) を計算した。VIF は以下の式で定義される。

$$VIF(j) = 1/(1 - R(j)^2) \quad (3)$$

ここで、 $R(j)^2$ は j 番目の説明変数を目的変数とし、他の説明変数を説明変数とした場合の重回帰分析における決定係数である。全ての VIF、つまり N1 においては 6 種類の、N2 においては 3 種類の VIF の値は 10 未満であったため、説明変数間には多重共線性はないとみなせる (Chatterjee and Hadi, 2012)。

以上より、ここで提案した重回帰モデルにより、ほぼ JND 以内の誤差範囲で N1 および N2 周波数を推定できると言える。

4.3. ナイーブな被験者の N1・N2 周波数推定精度

重回帰分析に含まれていない被験者の N1・N2 周波数を推定し、その精度を検証した。

被験者は正常な聴力をもつ男性 3 人 (OIS, TCY, MTZ) および女性 1 人 (CKT) であり、年齢は 21–25 歳である。実耳から計測した 6 箇所 (concha, incisura intertragica, concha, scapha, concha, pinna) の耳介形状パラメータを Table 4 に示す。これらと式(2)を用いて正面方向の N1・N2 周波数を推定した。また、実測 HRIR より N1・N2 を抽出し、その周波数を求めた。推定値、測定値、および誤差を Table 5 に示す。N1・N2 の推定誤差は 8 耳すべてにおいて JND 未満であった。比較的大きな誤差は被験者 CKT の左耳 (0.10 oct.)、MTZ の左耳 (0.09 oct.) および TCY の右耳 (0.09 oct.) で観察された。

Table 4 Six pinnae dimensions of four subjects (mm).

Subject	Ear	Width of		Length of		Depth of	Tilt of
		concha	incisura intertragica	concha	scapha	concha	pinna (°)
		x_2	x_3	x_6	x_8	x_d	x_a
OIS	L	19.2	7.9	23.3	19.5	12.9	28
	R	17.0	8.7	21.0	20.1	13.5	18
TCY	L	14.4	8.0	23.0	20.3	12.5	29
	R	14.3	7.6	22.8	20.4	12.9	23
CKT	L	15.7	7.4	19.7	18.3	13.9	22
	R	15.0	8.2	19.6	18.3	12.1	21
MTZ	L	17.6	7.6	22.6	17.2	11.2	28
	R	18.1	7.2	21.5	18.0	13.9	24

Table 5 Estimated and extracted frequencies of N1 and N2 and the residual errors for four subjects.

Subject	Ear	Estimated frequency [Hz]		Extracted frequency [Hz]		Residual error [oct.]	
		N1	N2	N1	N2	N1	N2
OIS	L	6749	9273	6938	9375	-0.04	-0.02
	R	7147	9779	6938	9281	0.04	0.08
TCY	L	6163	9249	6094	9656	0.02	-0.06
	R	6481	9221	6094	9375	0.09	-0.02
CKT	L	7358	10576	6844	10406	0.10	0.02
	R	7454	10920	7219	11250	0.05	-0.04
MTZ	L	7182	10364	6750	10875	0.09	-0.07
	R	7271	10061	7313	10125	-0.01	-0.01

5. best-matching HRTF の選出

HRTF データベースから best-matching HRTF を選出する方法を提案し、音像定位実験によりその有効性を検証した。ここで、best-matching HRTF とは推定値に最も近い周波数の N1・N2 をもつ HRTF と定義する。

5.1. 正中面内の best-matching HRTF の選出方法

HRTF データベースから best-matching HRTF を選出する物理指標として、NFD (notch frequency distance) を用いた。NFD は被験者 j の $HRTF_j$ と被験者 k の $HRTF_k$ の距離をオクターブスケールにより以下の式で表したものである (Iida and Ishii, 2011a)。

$$NFD_1 = \log_2\{f_{N1}(HRTF_j)/f_{N1}(HRTF_k)\} \text{ [oct.]} \quad (4)$$

$$NFD_2 = \log_2\{f_{N2}(HRTF_j)/f_{N2}(HRTF_k)\} \text{ [oct.]} \quad (5)$$

$$NFD = |NFD_1| + |NFD_2| \text{ [oct.]} \quad (6)$$

ここで、 f_{N1} および f_{N2} はそれぞれ N1 および N2 周波数を表す。

正面方向の N1・N2 周波数の推定値との NFD が最小となる HRTF をデータベースから選出し、best-matching HRTF とした (左耳と右耳では個別に選出)。加えて、上半球正中面の正面以外の方向の best-matching HRTF として、正面方向の best-matching HRTF の提供者の上半球正中面 HRTF を選出した。

今回用いたデータベースは日本人成人 120 耳の上半球正中面 7 方向の HRTF で構成されている。120 耳のうち 54 耳は 4 章の重回帰分析に用いたものである。なお、このデータベースは著者らの研究室のサイトで公開している (<http://www.iida-lab.it-chiba.ac.jp/e/>)。

5.2. best-matching HRTF の物理的精度

4 人の被験者 (OIS, TCY, CKT, MTZ) の正面方向の best-matching HRTF を N1・N2 周波数の推定値より選出した。

Fig.4 に 4 人の被験者本人の HRTF と best-matching HRTF のスペクトルを示す。best-matching HRTF の N1・N2 周波数 (●,▲) と被験者本人の HRTF (○,△) は近接している。また、ほとんど全ての耳において、best-matching HRTF (点線) と被験者本人の HRTF (実線) で同様のスペクトル構造が観察された。しかし、被験者 CKT の左耳の best-matching HRTF は本人の HRTF と類似しているとは言えない。best-matching HRTF の N1 は本人のそれと比較して浅く、N2 は深い。本方法ではノッチのレベルを考慮していないので、このようなことは起こり得る。

Table 6 に 正面方向の被験者本人の HRTF および best-matching HRTF の N1・N2 周波数を示す。残差は、耳介形状による N1・N2 周波数の推定誤差とデータベースから best-matching HRTF を選出する際の誤差の重

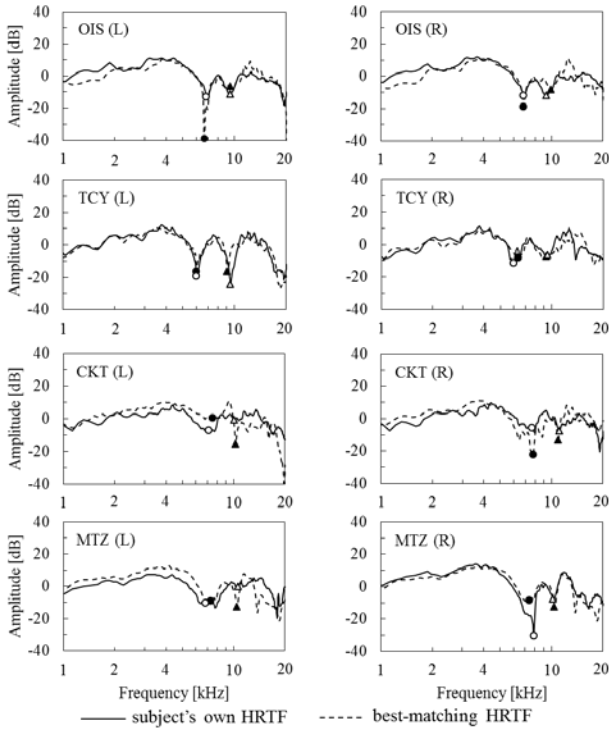


Figure 4 Amplitude spectrum of the subjects' own (solid line) and best-matching HRTFs (broken line) for the front direction. ○: N1 (own), ●: N1 (best-matching), △: N2 (own), ▲: N2 (best-matching).

Table 6 N1 and N2 frequencies of subjects' own HRTFs and best-matching HRTFs for the front direction.

Subject	Ear	best-matched frequency [Hz]		Extracted frequency [Hz]		Residual error [oct.]	
		N1	N2	N1	N2	N1	N2
OIS	L	6844	9375	6938	9375	-0.02	0.00
	R	6938	9844	6938	9281	0.00	0.08
TCY	L	6094	9188	6094	9656	0.00	-0.07
	R	6469	9188	6094	9375	0.09	-0.03
CKT	L	7406	10594	6844	10406	0.11	0.03
	R	7500	10875	7219	11250	0.06	-0.05
MTZ	L	7219	10313	6750	10875	0.10	-0.08
	R	7219	10313	7313	10125	-0.02	0.03

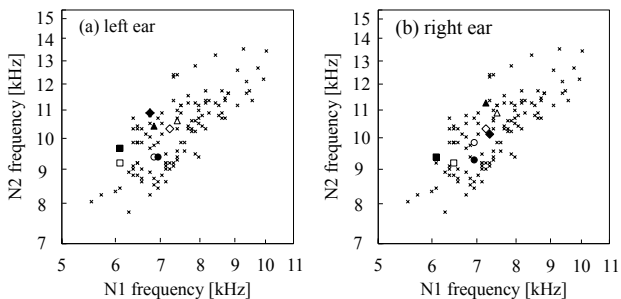


Figure 5 Scatter plot of the subjects' own HRTFs and best-matched HRTFs on the N1-N2 plane. (a): N1, (b): N2. ○: OIS (own), ●: OIS (best-matching), □: TCY (own), ■: TCY (best-matching), △: CKT (own), ▲: CKT (best-matching), ◇: MTZ (own), ◆: MTZ (best-matching), x: other 120 ears.

ね合わせである. N1・N2 の残差は 8 耳全てにおいて JND 未満であった. 比較的大きな誤差は被験者 CKT の左耳(0.11 oct.), MTZ の左耳(0.10 oct.)および TCY の右耳(0.09 oct.)で観察された. これは, Table 5 と同じ傾向であり, 推定誤差が支配的であると言える.

Fig.5 は 被験者本人の HRTF と best-matching HRTF の N1-N2 平面での散布図である. 120 耳の布置は広がっているが, いずれの被験者においても本人の HRTF と best-matching HRTF は近接している. ただし, 被験者 CKT と MTZ の左耳では, 他の耳と比較して距離が大きい.

次に, 上半球正中面 7 方向における残差を示す (Table 7). Table 6 で示したように 0°での残差は JND 未満であった. 他の 6 方向のほとんどにおいても残差は JND 未満であった. しかし, N1 については, 被験者 OIS の 180°の左耳(0.20 oct.), 150°の右耳(0.16 oct.), 被験者 CKT の 120°の右耳 (-0.16 oct.)において残差は JND を上回った.

以上より, 正方向だけでなく他の上半球正中面のほとんどの方向において, best-matching HRTFs は被験者本人の HRTF と類似のスペクトルを有していると考えられる.

Table 7 Residual errors in the N1 and N2 frequencies between the subjects' own HRTFs and best-matching HRTFs for each ear for seven vertical angles in the upper median plane (oct.).

Subject	Ear		Vertical angle (°)						
			0	30	60	90	120	150	180
OIS	L	N1	-0.02	0.10	-0.09	-0.03	-0.04	0.13	0.20
		N2	0.00	0.14	0.07	-0.04	0.06	0.11	0.08
	R	N1	-0.02	0.05	0.00	0.08	0.06	0.16	0.00
		N2	0.08	0.06	0.01	0.03	0.04	0.00	0.00
TCY	L	N1	-0.02	-0.04	-0.02	0.00	-0.07	0.01	-0.05
		N2	-0.09	0.01	-0.03	0.06	0.02	-0.02	0.08
	R	N1	0.09	0.00	0.00	-0.07	-0.05	-0.01	-0.03
		N2	-0.03	0.08	0.00	0.08	0.00	-0.10	0.05
CKT	L	N1	0.11	0.13	-0.03	-0.08	-0.09	-0.07	-0.09
		N2	0.03	-0.01	0.06	0.11	0.04	-0.07	-0.01
	R	N1	0.06	0.02	-0.12	-0.04	-0.16	-0.03	-0.06
		N2	-0.05	-0.08	-0.14	-0.03	-0.04	0.12	0.04
MTZ	L	N1	0.10	0.07	0.08	-0.06	-0.02	0.07	0.08
		N2	-0.08	0.08	-0.03	-0.07	-0.06	0.10	0.06
	R	N1	-0.02	0.09	0.00	-0.06	0.05	-0.09	-0.08
		N2	0.03	0.03	0.01	-0.01	-0.03	0.05	-0.05

5.3. best-matching HRTF の知覚的精度

5.3.1. 実験方法

A. HRTF を用いた音像定位実験

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), オープンタイプヘッドホン (AKG K1000), 耳栓型マイク, A/D コンバータ (Roland M-10MX)である.

Møller *et al.*(1995)は、外耳道入口を開放した状態における鼓膜での音圧は、閉塞した状態の外耳道入口の音圧に以下の補正 G を施すことによって得られることを示した。

$$G = \frac{1}{M \cdot PTF} \cdot \frac{Z_{\text{ear canal}} + Z_{\text{headphone}}}{Z_{\text{ear canal}} + Z_{\text{radiation}}} \quad (7)$$

ここで、 M は耳栓型マイクロホンの伝達関数、 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、 $Z_{\text{ear canal}}$ と $Z_{\text{headphone}}$ はそれぞれ外耳道およびヘッドホンのインピーダンス、 $Z_{\text{radiation}}$ 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、彼らは本実験で用いたヘッドホン (AKG K-1000) の PDR は 1 とみなせると報告している。

この原理に則り刺激を提示した。まず、マイクロホンの振動板が外耳道入口に位置するよう被験者に耳栓型マイクロホンを装着した。次に、被験者にヘッドホンを装着して 12 次の M 系列信号 (サンプリング周波数: 48kHz) を提示し、耳栓型マイクロホンまでの伝達関数 $M \cdot PTF$ を測定した。さらに、ヘッドホンを装着したまま耳栓型マイクロホンだけを取り外し、以下の式で作成した刺激 $P_{1,r}(\omega)$ をヘッドホンから提示した。

$$P_{1,r}(\omega) = S(\omega) \cdot HRTF_{1,r}(\omega) / (M_{1,r}(\omega) \cdot PTF_{1,r}(\omega)) \quad (8)$$

ここで、 $S(\omega)$, l , r はそれぞれ音源信号、左耳、右耳を表す。音源信号は 200Hz–17kHz の広帯域ホワイトノイズである。HRTF には、被験者本人の実測 HRTF と best-matching HRTF を用いた。

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。刺激の提示音圧レベルは、被験者の左右の外耳道入口で 63 dB SPL とした。両耳間時間差および両耳間レベル差はゼロに設定した。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。

被験者がターゲット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。被験者はノート PC の画面上の正中面を示す円周上に、音像方向をスタイラスペンでチェックした。頭内定位した場合はチェックボックスに印を入れた。刺激はランダムな順に提示し、各被験者は各刺激について 10 回ずつ回答した。被験者は OIS, TCY, CKT, MTZ の 4 人である。

B. 実音源を用いた音像定位実験

HRTF を用いた音像定位実験に先立って実音源による音像定位実験を行った。この実験の目的は被験者の正中面音像定位能力を確認することである。実験は無響室で行った。音源信号は 200Hz–17kHz の広帯域ホワイトノイズであり、上半球正中面の 30° 間隔の 7 方向に設置した小型スピーカ (FOSTEX FE83E) からラン

ダムな順に提示した。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。なお、中心周波数 250Hz から 16kHz までの 1/3 オクターブバンドレベルの偏差が 1dB 以内になるようイコライザで調整している。無響室は消灯し、回答に必要な手元灯のみ使用した。被験者はマッピング法により回答用紙に音像方向を記入した。各被験者は各刺激に対し 10 回ずつ回答した。

5.3.2. 実験結果

A. 実音源と各 HRTF に対する回答

Figs.6–9 に実音源、被験者本人の HRTF、best-matching HRTF 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、横軸は目標の上昇角である。円の直径は回答の頻度に比例している。なお、回答方向は 5° 間隔で丸めている。

被験者 OIS (Fig.6) では、実音源に対する回答 (a) は対角線を跨いで S 字カーブ状に分布している。本人の HRTF (b) に対しても実音源同様に S 字カーブを描いている。ただし、目標方向 60° および 120° に対する回答はやや上方にシフトしている。best-matching HRTFs (c) に対しては、 $N1 \cdot N2$ 周波数を推定した目標方向 0° では、本人 HRTF とほぼ同じ方向に知覚している。また、目標方向 30, 60, 180° での回答の分布も本人 HRTF と同様であった。一方、目標方向 90° および 120° では 120–150° に知覚する傾向があり、目標方向 150° での回答は上昇する傾向があった。

被験者 TCY (Fig.7) では、実音源に対する回答 (a) は目標方向 0–60° および 180° では対角線上に布置している。しかし、目標方向 90° および 120° では後方に、150° では上方に知覚する傾向がある。本人の HRTF (b) に対しては目標方向 120° および 150° では回答がばらつくが、他の方向では対角線上に布置している。best-matching HRTFs (c) に対しては、 $N1 \cdot N2$ 周波数を推定した目標方向 0° では、本人 HRTF とほぼ同じ方向に知覚している。30, 60, 180° でも目標方向付近に知覚している。目標方向 90° では回答のばらつきは本人の HRTF より大きい。実音源と同程度であった。120° では回答のばらつきは本人の HRTF と同程度であったが、上方から前方にシフトする傾向があった。150° では回答のばらつきは本人 HRTF よりも小さい。

被験者 CKT (Fig.8) では、実音源に対する回答 (a) は概ね対角線上に布置している。ただし、目標方向 0° において一度だけ上方に回答した。本人の HRTF (b) に対しては、概ね目標方向に知覚している。ただし、目標方向 150° では回答のばらつきは実音源よりも大きい。best-matching HRTFs (c) に対しては、 $N1 \cdot N2$ 周波数を推定した目標方向 0° では、本人 HRTF とほぼ同じ方向に知覚している。目標方向 150° および 180° でも本人 HRTF と同等の分布であった。30° では回答は上方にシ

フトする傾向がある．他の3方向(60, 90, 120°)では，回答のばらつきは本人 HRTF より大きい．

被験者 MTZ (Fig.9)では，実音源に対する回答(a)は対角線を跨いで S 字カーブ状に分布している．本人の HRTF(b)に対しても実音源同様に S 字カーブを描いている．ただし，目標方向 90°および 120°に対する回答はやや後方にシフトしている．best-matching HRTFs(c)に対しては，N1・N2 周波数を推定した目標方向 0°では，本人 HRTF とほぼ同じ方向に知覚している．目標方向 60, 120, 150, 180°においても回答の分布は本人 HRTF と同等であった．ただし，30°および 90°では回答のばらつきは本人 HRTF より大きい．

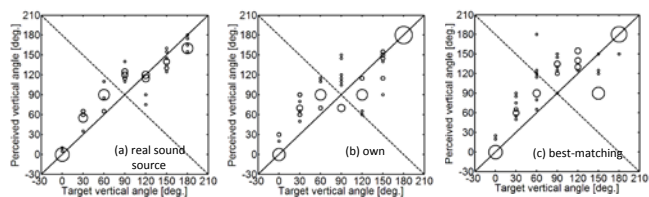


Figure 6 Responses of subject OIS to (a) the real sound sources, (b) the subject's own HRTFs, and (c) the best-matching HRTFs.

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.

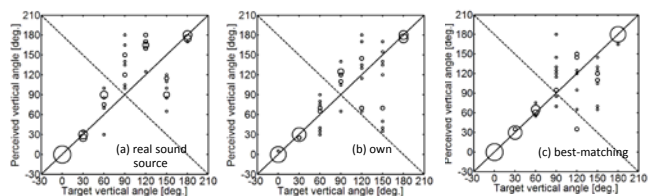


Figure 7 Responses of subject TCY to (a) the real sound sources, (b) the subject's own HRTFs, and (c) the best-matching HRTFs.

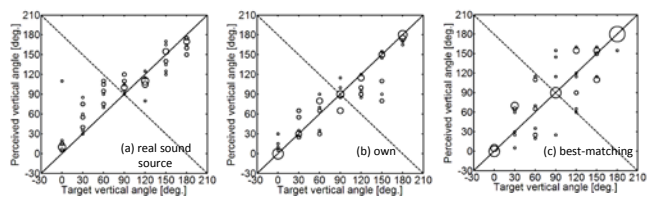


Figure 8 Responses of subject CKT to (a) the real sound sources, (b) the subject's own HRTFs, and (c) the best-matching HRTFs.

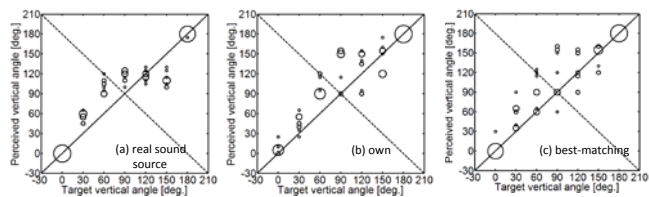


Figure 9 Responses of subject MTZ to (a) the real sound sources, (b) the subject's own HRTFs, and (c) the best-matching HRTFs.

B. 平均定位誤差

Table 8 に平均定位誤差を示す．ここで，平均定位誤差は目標の上昇角と知覚した上昇角の差の絶対値の平均値と定義する．

被験者 OIS では，best-matching HRTF の目標方向 0°および 180°の平均定位誤差は 5.2°および 3.4°であった．これらの値は本人の HRTF と同程度である．しかし，150°での平均定位誤差は 47.3°であり，本人 HRTF の約 3 倍となった．これは Table 7 に示した本人 HRTF と best-matching HRTF の N1 周波数の比較的大きな差異(右耳: 0.16 oct., 左耳: 0.13 oct.)に依るものと解釈できる．Table 7 にはもう 1 つの大きな N1 周波数の差異(180°における左耳: 0.20 oct.)が示されているが，この場合は右耳での差異が N1・N2 とともに 0.00 oct.であり，これが小さな平均定位誤差(3.4°)に貢献していると考えられる．

被験者 TCY では，best-matching HRTF の目標方向 0°および 180°の平均定位誤差は 0.3°および 2.4°であった．これらの値は本人の HRTF と同程度である．また，他の 5 方向においても best-matching HRTF の平均定位誤差は本人 HRTF のそれと同等であった．

被験者 CKT では，best-matching HRTF の目標方向 0°および 180°の平均定位誤差は 2.1°および 3.5°であった．これらの値は本人の HRTF と同程度である．しかし，30°での平均定位誤差は 29.2°であり，本人 HRTF の約 2 倍となった．これは Table 7 に示した本人 HRTF と best-matching HRTF の N1 周波数の比較的大きな差異(左耳: 0.13 oct.)に依るものと解釈できる．また，120°での平均定位誤差は 31.9°であり，本人 HRTF の約 2 倍となった．これも本人 HRTF と best-matching HRTF の N1 周波数の比較的大きな差異(右耳: -0.16 oct.)に依るものと解釈できる．

被験者 MYZ では，best-matching HRTF の目標方向 0°および 180°の平均定位誤差は 3.3°および 0.5°であった．これらの値は本人の HRTF と同程度である．また，他の 5 方向においても best-matching HRTF の平均定位誤差は本人 HRTF のそれと同等であった

Table 8 Mean localization errors for each subject, HRTF, and target vertical angle (°).

Subject	HRTF	Target vertical angle (°)							
		0	30	60	90	120	150	180	Ave.
OIS	real sound source	2.5	25.9	27.4	32.0	9.5	12.4	17.5	18.2
	own HRTF	8.7	40.4	30.5	28.1	29.7	16.4	0.5	22.0
	best-matching HRTF	5.2	36.7	44.3	39.8	20.4	47.3	3.4	28.2
TCY	real sound source	0.4	2.8	25.7	45.7	44.0	52.2	2.8	24.8
	own HRTF	0.7	1.8	17.3	29.5	34.2	56.8	3.0	20.5
	best-matching HRTF	0.3	1.9	4.9	30.5	36.7	40.0	2.4	16.7
CKT	real sound source	20.8	25.2	32.9	14.8	13.3	13.2	15.4	19.4
	own HRTF	5.9	13.6	19.1	12.1	17.4	29.0	4.1	14.4
	best-matching HRTF	2.1	29.2	30.8	21.3	31.9	17.3	3.5	19.5
MTZ	real sound source	0.9	25.1	41.6	26.9	5.4	39.0	0.8	20.0
	own HRTF	5.9	16.6	39.2	46.2	25.4	17.1	1.1	21.6
	best-matching HRTF	3.3	24.8	27.9	38.5	20.6	13.1	0.5	18.4

C. 前後誤判定率

Table 9 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。

全ての被験者において、best-matching HRTF の目標方向 0°および 180°の前後誤判定率は本人 HRTF 同様 0%であった。他の 5 方向についても best-matching HRTF の前後誤判定率は本人 HRTF と同等であった。ただし、被験者 OIS の 150°, CKT の 60°, MTZ の 120°では best-matching HRTF の前後誤判定率は本人 HRTF よりも高い。

Table 9 Ratio of front-back confusion for each subject, HRTF, and target vertical angle (%).

Subject	HRTF	Target vertical angle (°)						Ave.	
		0	30	60	90	120	150		180
OIS	real sound source	0	0	40	-	10	0	0	8.3
	own HRTF	0	20	70	-	20	10	0	20.0
	best-matching HRTF	0	10	60	-	0	30	0	16.7
TCY	real sound source	0	0	10	-	0	20	0	5.0
	own HRTF	0	0	10	-	30	50	0	15.0
	best-matching HRTF	0	0	0	-	30	20	0	8.3
CKT	real sound source	10	0	60	-	10	0	0	13.3
	own HRTF	0	0	10	-	20	30	0	10.0
	best-matching HRTF	0	0	30	-	20	0	0	8.3
MTZ	real sound source	0	0	100	-	0	0	0	16.7
	own HRTF	0	0	100	-	0	0	0	16.7
	best-matching HRTF	0	10	50	-	20	0	0	13.3

D. 頭内定位率

全ての被験者において、best-matching HRTF および本人 HRTF では頭内定位は生じなかった。

E. 音像定位実験結果のまとめ

上記の結果は、受聴者の N1・N2 周波数を推定した正面方向(0°)において、best-matching HRTF は本人 HRTF と同等精度の音像定位を実現できることを示している。また、best-matching HRTF は後方(180°)においても、前方と同様の音像定位精度を有していることを示している。他の 5 方向については、被験者の一部で本人 HRTF よりも音像定位精度は低下した。

6. 考察

6.1. 従来の HRTF 個人化方法との比較

本提案法では、受聴者の 6 箇所耳介形状を計測し、重回帰式を用いて N1・N2 周波数を推定し、それに最も近い best-matching HRTF をデータベースから選出した。このプロセスに必要な時間は約 3 分である。これは従来法より際立って速い。Middlebrooks *et al.* (2000) は受聴者に適した scale factor を見いだすには 1 ブロックで 20 分要する聴覚実験が 1 から 3 ブロック必要であると述べている。Iwaya (2006) はトーナメント方式の聴覚実験を提案しているが、32 の HRTF セットから受聴者に適したセットを選出するのに 15 分要すると報

告している。この場合、データベースに含まれる HRTF が増えれば増えるほど聴覚実験が必要とする時間が増えることに留意しなければならない。

また、本提案法は作業者に特別な訓練も装置も要求しないという利点を有する。

6.2. 3次元空間の任意の方向への展開

本研究では、受聴者に適した正面方向の HRTF の選出方法を提案し、上半球正中面の他の方向の HRTF については、正面方向の HRTF の提供者のものをを用いた。

ここでさらに、3次元空間の任意の方向への精度の高い音像制御について考察する。前述の HRTF データベースに必要な方向数の HRTF が用意されていれば問題はないが、それは理論的には膨大な数になる(例えば正面では方向知覚の弁別閾は約 1°である)。したがって、ヒトが実音源に対して知覚するのと厳密な意味で同精度の 3次元音像再生を実現することは原理的に不可能である。

そこで、できるだけ少数の方向の HRTF で 3次元空間の任意の方向に近似的に音像を制御することを考える。Morimoto and Aokata (1984) は、受聴者は音像の側方角と上昇角を独立に知覚し、側方角知覚の手掛かりは両耳間差であり、上昇角知覚の手掛かりはスペクトルであると報告した。さらに Morimoto *et al.* (2003) は 3次元空間の任意の方向への音像制御は正中面の HRTF と両耳間時間差を組み合わせることで実現できることを示した。これらの知見は、正中面内の best-matching HRTF と両耳間時間差を組み合わせることで任意の方向への音像制御が実現できる可能性を示唆しているが、これについては今後の課題としたい。

7. おわりに

精度の高い音像制御を実現することを目的として、受聴者の耳介形状から HRTF の N1・N2 周波数を推定する方法を提案した。さらに、推定値に最も近い N1・N2 周波数をもつ best-matching HRTF を HRTF データベースから選出する方法を提案した。これらの提案法の有効性を検証するために音像定位実験を行った。結論を以下にまとめる。

(1) 日本人成人 54 耳の正面方向の N1・N2・P1 周波数を実測 HRTF から抽出した。個人差は、それぞれ 0.74, 0.71, 0.31 オクターブであった。

(2) 54 耳の正面方向の N1・N2 周波数を目的変数、6 箇所耳介形状を説明変数として重回帰分析を行った。N1・N2 の重相関係数はそれぞれ 0.81 および 0.82 であった。

(3) 4 人のナイーブな被験者の正面方向の N1・N2 周波数を 6 箇所耳介形状から推定した。8 耳全てにおいて残差は JND 未満であった。

(4) この4人の被験者を用いて上半球正中面で音像定位実験を行った。その結果、N1・N2を推定した0°(正面)では、best-matching HRTFの音像定位精度は被験者本人のHRTFと同等であった。

(5) 180°(後方)においても、best-matching HRTFの音像定位精度は被験者本人のHRTFと同等であった。

(6) 他の上方の5方向では、幾人かの被験者の音像定位精度は本人のHRTFよりも低下した。

(7) 提案した個人化プロセスの所要時間は約3分で、従来の方法よりも顕著に短い。また、耳介形状の計測は特別な訓練や設備を必要としない。

謝 辞

本研究の一部は文部科学省私立大学戦略的研究基盤形成支援事業(S1311003)の支援を受けた。また、神戸大学の森本政之名誉教授には有益な助言をいただいた。千葉工業大学大学院生の岡松俊哉君、坂口慎治君、土屋宏樹君には耳介形状の測定に協力いただいた。記して感謝いたします。

文 献

- Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C. (2001). "The CIPIC HRTF database," Proc. IEEE Workshop on Application of signal processing to audio and acoustics, 99–102.
- Butler, A. and Belendiuk, K. (1977). "Spectral cues utilizes in the localization of sound in the median sagittal plane," J. Acoust. Soc. Am. **61**, 1264–1269.
- Chatterjee, S. and Hadi, A. S. (2012), *Regression Analysis by Example*, 5th edn. New Jersey, John Wiley & Sons.
- 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.
- Hu, H., Chen, L., and Wu, Z. (2008). "The estimation of personalized HRTFs in individual VAS," Proc. Fourth International Conference on Natural Computation, 203–207.
- Hugeng, W. W. and Gunawan, D. (2010). "Improved method for individualization of head-related transfer functions on horizontal plane using reduced number of anthropometric measurements," J. Telecommun. **2**, 31–41.
- 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," Applied Acoustics **68**, 835–850.
- Iida, K. and Ishii, Y. (2011a). "3D sound image control by individualized parametric head-related transfer functions," Proc. Inter-Noise 2011, 428959, Osaka, Japan.
- Iida, K. and Ishii, Y. (2011b). "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.
- Iwaya, Y. (2006). "Individualization of head-related transfer functions with tournament-style listening test: Listening with other's ears," Acoust. Sci. & Tech. **27**, 340–343.
- Kahana, Y. and Nelson, P. A. (2006). "Numerical modeling of the spatial acoustic response of the human pinna," J. Sound Vib. **292**, 148–178.
- Katz, B. F. (2001). "Boundary element method calculation of individual head-related transfer function I Rigid model calculation," J. Acoust. Soc. Am. **110**, 2440–2248.
- Kistler, D. J. and Wightman, F. L. (1992). "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," J. Acoust. Soc. Am. **91**, 1637–1647.
- Kreuzer, W., Majdak, P., and Chen, Z. (2009). "Fast multipole boundary element method to calculate head-related transfer functions for a wide frequency range," J. Acoust. Soc. Am. **126**, 1280–1290.
- 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.
- Mehrgardt, S. and Mellert, V. (1977). "Transformation characteristics of the external human ear," J. Acoust. Soc. Am. **61**, 1567–1576.
- Middlebrooks, J. C. and Green, D. M. (1992). "Observations on a principal components analysis of head-related transfer functions," J. Acoust. Soc. Am. **92**, 597–599.
- Middlebrooks, J. C. (1999a). "Individual differences in external-ear transfer functions reduced by scaling in frequency," J. Acoust. Soc. Am. **106**, 1480–1492.
- Middlebrooks, J. C. (1999b). "Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency," J. Acoust. Soc. Am. **106**, 1493–1510.
- Middlebrooks, J. C., Macpherson, E. A., and Onsan, Z. A. (2000). "Psychophysical customization of directional transfer functions for virtual sound localization," J. Acoust. Soc. Am. **108**, 3088–3091.
- 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.
- Morimoto, M. and Ando, Y. (1980). "On the simulation of sound localization," J. Acoust. Soc. Jpn. (E) **1**, 167–174.
- Morimoto, M. and Aokata, H. (1984). "Localization cues of sound sources in the upper hemisphere," J. Acoust. Soc. Jpn. (E) **5**, 165–173.
- Morimoto, M., Iida, K., and Itoh, M. (2003). "Upper hemisphere sound localization using head-related transfer functions in the median plane and interaural differences," Acoust. Sci. & Tech. **24**, 267–275.
- Musicant, A. and Butler, R. (1984). "The influence of pinnae-based spectral cues on sound localization," J.

Acoust. Soc. Am. **75**, 1195–1200.

Rodriguez, S. G. and Ramirez, M. A. (2005). “Extracting and modeling approximated pinna-related transfer functions from HRTF data,” Proc. Eleventh Meeting of the International Conference on Auditory Display, 269–273, Limerick, Ireland.

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.

Sottek, R. and Genuit, K. (1999). “Physical modeling of individual head-related transfer functions,” J. Acoust. Soc. Am. **105**, Issue 2, 1162.

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.

Xu, S., Li, Z., and Salvendy, G. (2008). “Improved method to individualize head-related transfer function using anthropometric measurements,” Acoust. Sci. & Tech. **29**, 388–390.

Zotkin, D. N., Hwang, J., Duraiswami, R., and Davis, L. S. (2003). “HRTF personalization using anthropometric measurements,” Proc. IEEE workshop on applications of signal processing to audio and acoustics, 157–160, New Paltz, NY, USA.

飯田一博, 石井要次, 西岡伸介 (2013), “耳介形状から推定したスペクトラルノッチ周波数に基づいた頭部伝達関数の個人化,” 日本音響学会聴覚研究会資料, H-2013-91.

西岡伸介, 石井要次, 飯田一博 (2013年9月), “上昇角知覚に関する頭部伝達関数の第1ピーク周波数の弁別閾,” 日本音響学会講演論文集, 859-860.