

## 受聴者の耳介形状による 頭部伝達関数のスペクトラルノッチレベルの推定

☆海老沼 佑輔 (千葉工大院・工学研), 飯田 一博 (千葉工大)

### 1 はじめに

受聴者本人の頭部伝達関数(HRTF)を鼓膜上で再現することで任意の方向に音像を制御できるが, 他人の HRTF を用いると, 前後誤判定や頭内定位などの現象がしばしば発生する[1].

しかしながら, すべての受聴者について HRTF を測定することは, 膨大な時間と無響室などの設備が必要であり, 現実的ではない. 一方, 耳介の写真などから受聴者の HRTF を推定することが考えられるが[2], 受聴者の HRTF がもつ全ての情報を推定することは困難である. そこで, ヒトが方向知覚の手掛かりとして利用している情報のみを推定することを考える.

方向知覚の手掛かりのうち, 左右方向の両耳間差キューについては, 両耳間距離などの身体的特徴から予測することができる[3].

前後・上下方向のスペクトラルキューについては, HRTF の 4kHz 付近に存在するピーク(P1), および 4kHz 以上に存在する最初のノッチ(N1)とその次のノッチ(N2)(Fig. 1)が知られている[4].

Iida *et al.* [5]は, 耳介形状パラメータから, 受聴者の N1, N2, P1 周波数を推定し, HRTF データベースから N1, N2 周波数が近いものを選出することにより, 精度の高い3次元音像制御ができることを示した. しかし, レベルを考慮していないため, データベースから選出された HRTF の N1, N2 レベルは本人の N1, N2 レベルに近いとは限らない.

本研究では, N1, N2, P1 レベルを推定する方法を検討する.

### 2 分析に用いたデータ

#### 2.1 N1, N2, P1 レベル

28 名の成人の上半球正中面(30° 間隔)の HRTF を用いた.

N1, N2 の抽出には Iida *et al.* [5]が提案した

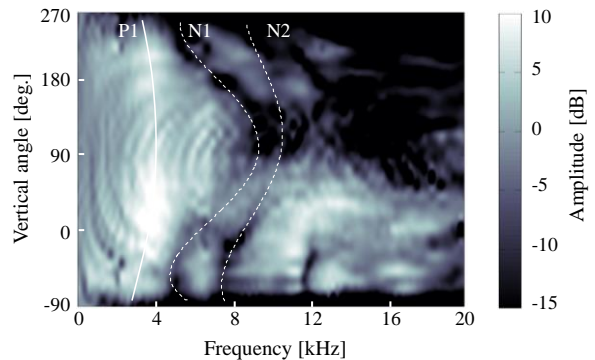


Fig. 1 Amplitude spectrum of HRTF in the median plane.

方法を用いた. これは頭部インパルス応答の初期応答部分(1 ms)を時間窓で切り出し, それをフーリエ変換し, 最も周波数の低い極大値を P1, それよりも高い帯域で最も周波数の低い 2 つの極小値を N1, N2 とする方法である. 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. ここで, ノッチレベルは, 初期応答から抽出した N1, N2, P1 と周波数が最も近い全応答長の HRTF のピーク, ノッチのレベルとした(Fig. 2). 54 耳の正面方向の N1, N2, P1 レベルの分布を Fig. 3 に示す. N1 レベルは -44.8 ~ 0.5 dB, N2 レベルは -33.1 ~ 2.5 dB, P1 レベルは, 5.1 ~ 18.4 dB の範囲に分布した.

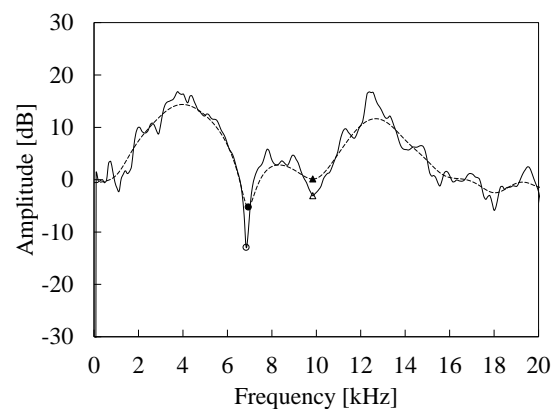


Fig. 2 HRTFs calculated from early part (1ms) of hrir (solid line) and full length of hrir (broken line).

\*Estimation of spectral notch levels of listener's head-related transfer function from anthropometry of the listener's pinna, by EBINUMA Yusuke and IIDA Kazuhiro, (Chiba Institute of Technology).

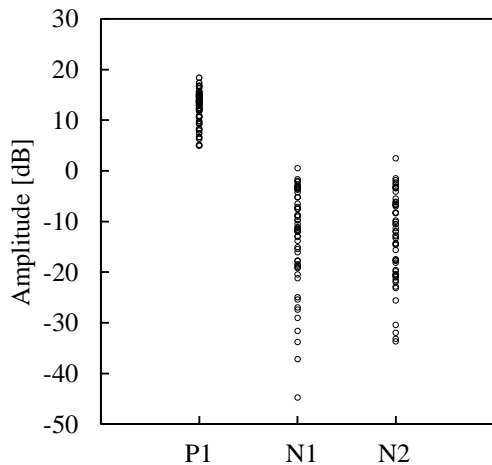


Fig. 3 Distribution of N1, N2 and P1 levels for the front direction for 54 ears.

## 2.2 耳介形状

N1, N2, P1 を抽出した 54 耳の 10 種類の耳介形状パラメータ (Fig.4) の計測値を Fig. 5 に示す.  $x_1 \sim x_8$  および  $x_d$  の分布範囲は, 10 ~ 20 mm であり, 耳介の傾き  $x_a$  は  $4 \sim 40^\circ$  の広い範囲に分布している.

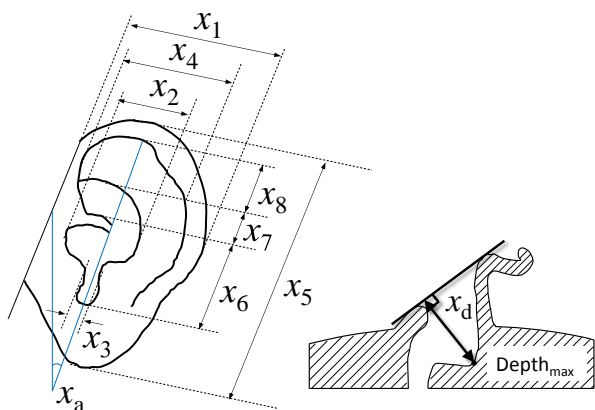


Fig. 4 Ten anthropometric parameters of the pinna.

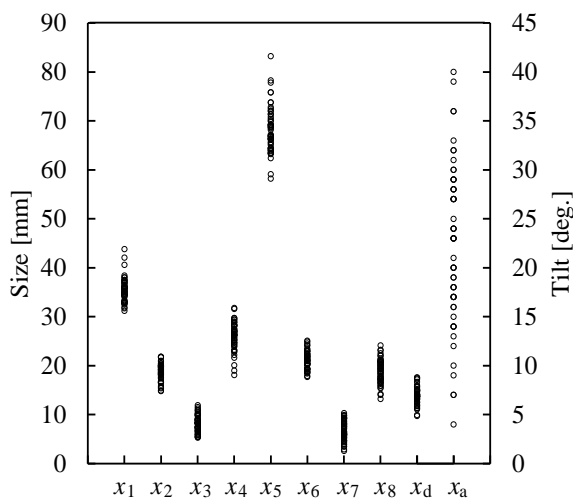


Fig. 5 Distribution of 10 measured anthropometric parameters for 54 ears.

## 3 正面方向の N1, N2, P1 レベルの推定

54 耳の正面方向の N1, N2, P1 レベルを目的変数, 耳介形状パラメータの全てを説明変数とした重回帰分析 (式 1) を行った. これを重回帰モデル A とする.

$$f(S)_{N1|N2|P1} = a_1x_1 + a_2x_2 + \dots + a_nx_n + b \text{ [dB]} \quad (1)$$

ここで,  $S$ ,  $a_i$ ,  $b$ ,  $x_i$  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.

得られた統計値を Table 1 に示す. また, 実測値と推定値の関係を Fig.6 に示す. N1, N2, P1 の重相関係数は, それぞれ 0.39, 0.58, 0.64 であった. しかし, 耳介形状パラメータの  $p$  値については, N2 の  $x_8, x_d$  を除き 0.05 を超えた. そこで, 目的変数を N1, N2, P1 とした場合に, 説明変数の個数を変えながらすべての説明変数の組み合わせについて重回帰分析を行い, すべての説明変数の  $p$  値が 0.05 未満か

Table 1 Statistics of multiple regression model A for the front direction.

	Correlation coefficient	Significance level	Absolute mean residual error [dB]
N1	0.39	0.6752	7.4
N2	0.58	0.0431	5.6
P1	0.64	0.0070	1.9

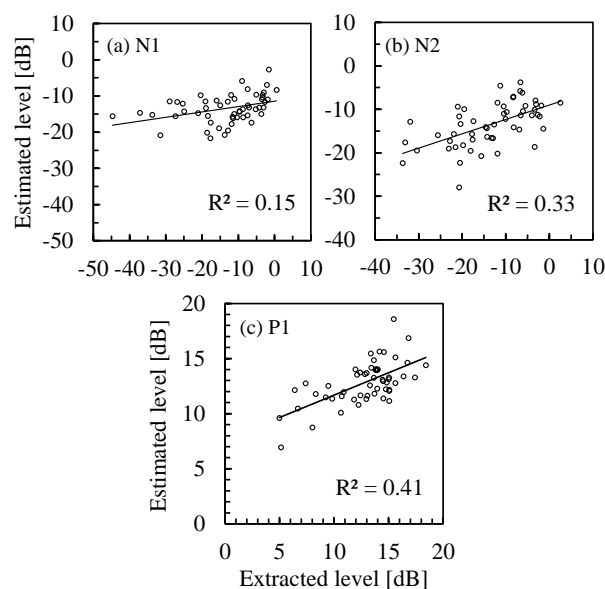


Fig. 6 Relationship between the levels extracted from the measured HRIR and the levels estimated from the listener's anthropometric parameters using the multiple regression model A. (a)N1; (b)N2; (c)P1.  $R^2$  denotes the coefficient of determination.

つ重回帰係数が最大となるの組み合わせを求めた。

その結果, N2 については 2 種類( $x_8, x_d$ ), P1 については 3 種類( $x_4, x_5, x_8$ )の耳介形状パラメータが選ばれた. これを回帰モデル B とする. また, N1 は, いずれの耳介形状パラメータにおいても p 値が 0.05 未満となる組み合わせは存在しなかった. つまり, 耳介形状パラメータの中で, N2 においては耳介の窪みの長さ, 深さが寄与し, P1 においては耳介の窪みの幅, 長さが寄与していると言える. 目的変数を N2, P1, 説明変数を耳介形状パラメータの 2 種類( $x_8, x_d$ ), 3 種類( $x_4, x_5, x_8$ )とした重回帰分析の統計値, および回帰係数  $a_i$ , p 値を Table 2, 3 に示す. また, 実測値と推定値の関係を Fig. 7 に示す. なお, 変数間での多重共線性はみられなかった.

N2, P1 の重相関係数はそれぞれ 0.51, 0.55 であった. 平均絶対残差はそれぞれ 5.9, 2.0 dB であった.

石井ら[6]は N1, N2, P1 で構成した pHRTF を用いて, 周波数を受聴者本人の N1, N2 の周波数で固定し, N1, N2 レベルをどちらも -20dB 以下の pHRTF を提示した場合, 音像定位精度が比較的に高い結果となることを示し

Table 2 Statistics of multiple regression model B for the front direction.

	Correlation coefficient	Significance level	Absolute mean residual error [dB]
N2	0.51	0.0005	5.9
P1	0.55	0.0004	2.0

Table 3 Multiple regression coefficients and p-values for the front direction of the multiple regression model B.

	Regression coefficient		p-value	
	N2	P1	N2	P1
$a_1$				
$a_2$				
$a_3$				
$a_4$		-0.54		0.02
$a_5$		-0.34		0.02
$a_6$				
$a_7$				
$a_8$	1.26	0.52	0.01	3.E-03
$a_d$	1.85		3.E-03	
$a_a$				
b	-62.98	16.44	2.E-06	1.E-03

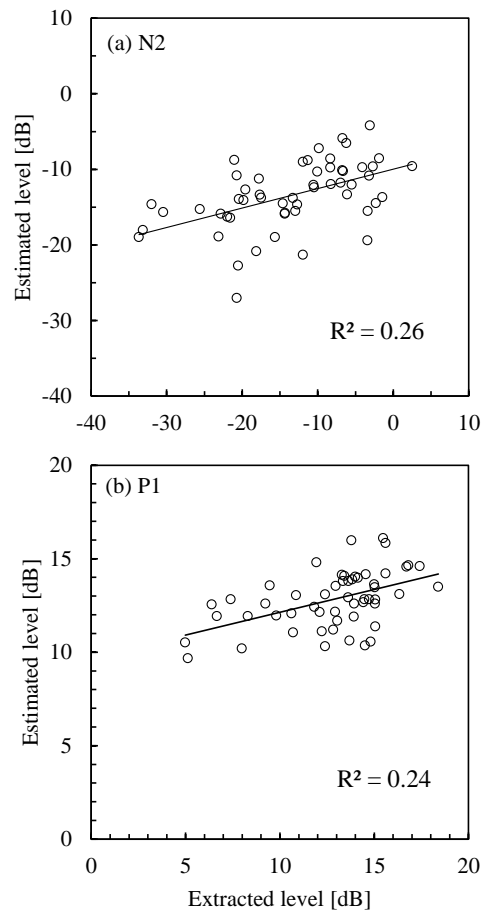


Fig. 7 Relationship between the levels extracted from the measured HRIR and the levels estimated from the listener's anthropometric parameters using the multiple regression model B. (a)N2; (b)P1.  $R^2$  denotes the coefficient of determination.

た. つまり, -20 dB 以下のレベルの違いが音像定位精度に及ぼす影響は小さい. そこで, -20 dB 以上の N1, N2 について, そのレベルの推定を試みる.

54 耳の中から N1, N2 レベルがそれぞれ -20 dB 以上の, 43 耳, 40 耳を用いて, 重回帰分析を行った. その結果, N1 については 6 種類( $x_1, x_2, x_3, x_4, x_5, x_8$ ), N2 については 2 種類( $x_8, x_d$ )の耳介形状パラメータが選ばれた. これを回帰モデル C とする. つまり, 耳介形状パラメータの中では, N1 においては, 耳介の窪みの幅, 長さ, 深さが寄与し, 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. また, N2 においては, 54 耳を用いた重回帰分析で選択された耳介形状パラメータと同様のものが選ばれた. 目的変数を N1, N2, 説明変数を耳介形状パラメータの 6 種類( $x_1, x_2, x_3, x_4, x_5, x_8$ ), 2 種類( $x_8, x_d$ )とし

Table 4 Statistics of multiple regression model C for the front direction.

	Correlation coefficient	Significance level	Absolute mean residual error [dB]
N1	0.66	0.0014	3.5
N2	0.43	0.0227	4.1

Table 5 Multiple regression coefficients and p-values for the front direction of the multiple regression model C.

	Regression coefficient		p-value	
	N1	N2	N1	N2
$a_1$	1.18		0.01	
$a_2$	1.46		0.01	
$a_3$	-1.70		3.E-03	
$a_4$	-1.33		1.E-03	
$a_5$	-0.39		0.04	
$a_6$				
$a_7$				
$a_8$	1.04	1.03	0.01	0.04
$a_d$		0.86		0.02
$a_a$				
b	-22.60	-47.83	0.13	1.E-03

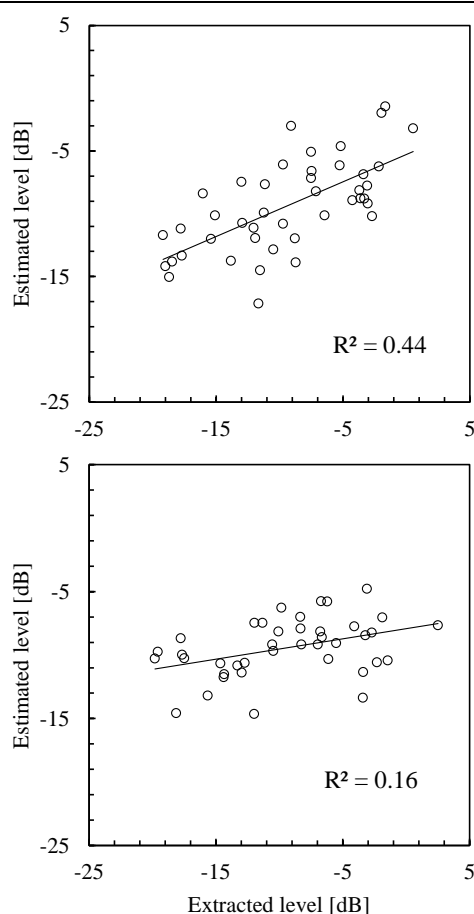


Fig. 8 Relationship between the levels extracted from the measured HRIR and the levels estimated from the listener's anthropometric parameters using the multiple regression model C. (a)N1; (b)N2.  $R^2$  denotes the coefficient of determination.

た重回帰分析の統計値, および回帰係数  $a_i$ ,  $p$  値を Table 4, 5 に示す. また, 実測値と推定値の関係を Fig.8 に示す. なお, 変数間での多重共線性はみられなかった.

N1, N2 の重相関係数はそれぞれ 0.66, 0.43 であった. 平均絶対残差はそれぞれ 3.5, 4.1 dB であった. N1 においては, 回帰モデル B では説明変数の  $p$  値が 0.05 未満となる回帰式を求めることができなかったが, 回帰モデル C では説明変数の  $p$  値が 0.05 未満となる回帰式を求めることができた. N2 においては, 回帰モデル B と比較し, 重相関係数, 絶対平均残差は共に小さくなった.

#### 4 おわりに

正面方向の N1, N2, P1 レベルを目的変数, 耳介形状パラメータを説明変数とした重回帰分析を行った. その結果, 以下のことを示した.

- 1) 54 耳による重相関分析の結果, N2, P1 レベルの実測値と予測値の間の重相関係数は, それぞれ 0.51, 0.55 であった. N1 レベルについては, 説明変数の  $p$  値が 0.05 未満となる回帰モデルが存在しなかった.
- 2) N1, N2 レベルが -20 dB 以上の 40 耳(N1), 43 耳(N2)を用いて, 重回帰分析を行った. その結果, N1, N2 レベルの実測値と予測値の間の重相関係数は, それぞれ 0.63, 0.43 であった.

#### 謝辞

本研究の一部は文科省私立大学戦略的研究基盤形成支援事業(S1311003)により実施した.

#### 参考文献

- [1] M. Morimoto and Y. Ando, J. Acoust. Soc. Jpn. (E) 1, 167-174, 1980.
- [2] R. Sottek and K. Genuit, Proc. DAGA, 1999
- [3] Algazi, J. Audio Eng. Soc., vol.49, pp. 472-479, 2001
- [4] K. Iida *et al.*, J. Acoust. Soc. Am, vol.68, pp. 835-850, 2007
- [5] K.Iida *et al.*, J. Acoust. Soc. Am, 136(1), pp. 317-333, 2014.
- [6] 石井, 飯田, 音講論(春), pp. 581-584, 2010.