

聴覚の仰角知覚メカニズムに基づいた音源方向推定*

○飯田一博 (千葉工大)

1 はじめに

著者は、HRTF の振幅スペクトルを複数の Peak と Notch に分解し、それらの一部または全部を中心周波数、レベル、尖鋭度で表現して再構成する Parametric HRTF(Fig.1)を提案し、正中面内の音源定位実験で第 1 Notch (N1)、第 2 Notch (N2)だけを用いた Parametric HRTF が実測 HRTF と同等の定位精度が得られること(Fig.2)を明らかにした。加えて、2つの Notch 周波数が音源の仰角に強く依存していることより、N1、N2 が仰角知覚のスペクトラルキューであると結論付けた[1]。

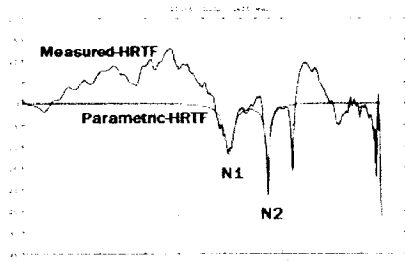


Fig.1 An example of measured HRTF (blue line) and Parametric HRTF (green line).

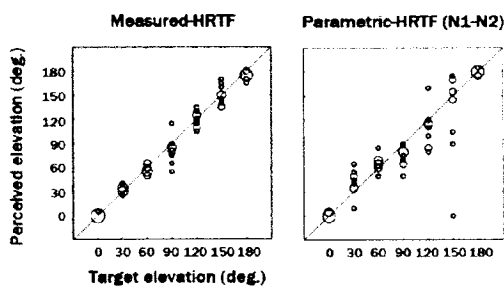


Fig.2 Responses to stimuli of measured and Parametric-HRTF (N1-N2) in the median plane.

これらの知見を利用すれば、耳入力信号から音源の仰角を推定できる可能性がある。音源方向推定については、これまでに様々な研究があるが[2]、本稿では仰角知覚メカニズムに基づいて、耳入力信号からスペクトラルキューを抽出し、仰角を推定することを試みた。

2 仰角推定方法

2.1 基本的な方針

本研究では、以下のような考え方で推定方法を検討した。

A) 入力情報について：

両耳入力信号のみを用いる。これは、バイノーラルヒアリングと入力信号を共有できることを意図している。

B) 信号処理アルゴリズムについて：

聴覚メカニズムとして既知の処理だけを用いる。機械的操作を加えて徒に推定精度を向上させることよりも、聴覚処理としての妥当性と応用展開の汎用性を重視する。本研究で用いた聴覚処理の知見は、具体的には以下のとおりである。

- ・ 仰角知覚は単耳スペクトル情報に基づいており、左右の耳で個別に処理されている[3]
- ・ 仰角知覚は音源信号の種類に関する *a priori* な知識を必要としない[4]
- ・ 仰角知覚は N1, N2 の周波数を手掛かりとしている[1]

2.2 仰角推定アルゴリズム

上記の方針にしたがって以下のような手順で信号処理を行う。

- ・ 両耳入力信号をそれぞれ周波数領域へ変換
- ・ 単耳振幅スペクトルを算出
- ・ 移動平均によりスペクトル包絡線を抽出
- ・ Notch 周波数を検出
- ・ N1-N2 データベースとの照合により仰角算出

N1-N2 データベースは、HRTF から抽出した N1、N2 の周波数を音源の仰角の関数として表したものである。Fig.3 は、上半球正中面における N1、N2 の周波数を示している。これを見ると、N1、N2 の周波数は仰角の増加に対して単調に変化するのではなく、上に凸

* A study on estimation of sound source elevation based on the vertical sound localization mechanism by IIDA, Kazuhiro (Chiba Institute of Technology)

の関数となっている（1つの Notch だけでは仰角知覚ができないことは、このような非単調性でも説明できる）。そこで、4次関数を用いて近似を行った。決定係数 R^2 はそれぞれ 0.98(N1), 0.99(N2)であった。

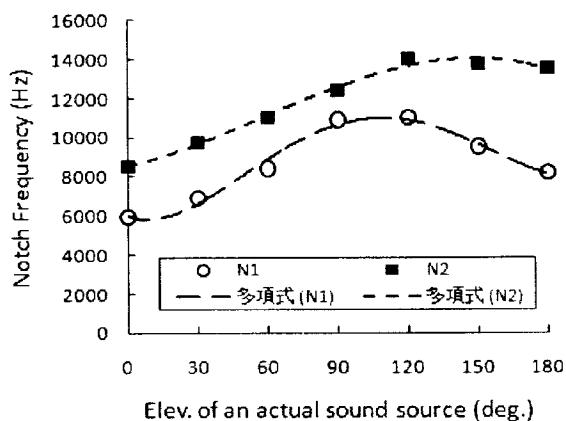


Fig.3 Relation between elevation of a sound source and frequencies of N1 and N2.

3 仰角推定シミュレーション

仰角推定方法の有効性を検証するためにシミュレーションを行った。音源信号はホワイトノイズ、ピンクノイズ、男声アナウンス、女声アナウンス、音楽(ポップス)で、継続時間は1sである。各音源に正中面内の7方向(0-180°, 30°間隔)の HRTF のいずれかを畳込んで耳入力信号とした。つまり、今回用いた耳入力信号は、音源はただ1つであり、しかも無響室条件である。まずは、このような理想的な条件でシミュレーションを行った。サンプリング周波数は48kHzである。

Fig.4 に耳入力信号から N1, N2 を抽出する過程の一例を示す。(a)は仰角 0°の場合の HRTF で、(b)は(a)に女声アナウンスを畳んだ信号の振幅スペクトルである(縦軸のスケールが異なるので注意)。(c)はスペクトル包絡線とそこから抽出した N1, N2 を示す。このようにして得られた N1, N2 の周波数と Fig.3 の近似式で得られる周波数の差が最も小さくなる仰角を求め、それを推定値とした。

Fig.5 にシミュレーション結果を示す。音源の種類に関わらず、概ね正しく仰角を推定している。ただし、女声アナウンスの 30°および音楽の 0°の場合に後方に誤って推定している。N1, N2 の振舞いが前方と後方で似て

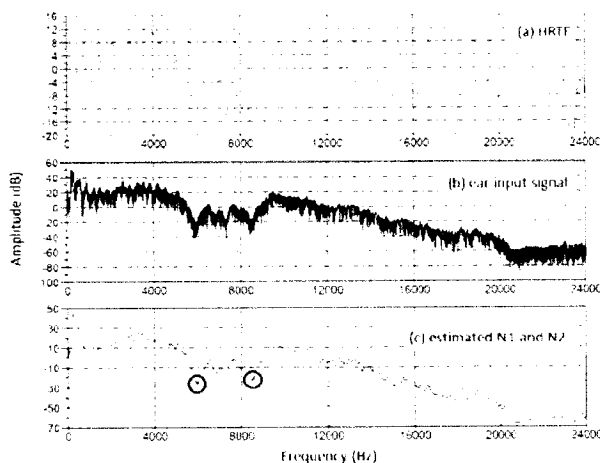


Fig.4 An example of detection process of N1 and N2 from ear-input signal. ○:detected N1 and N2.

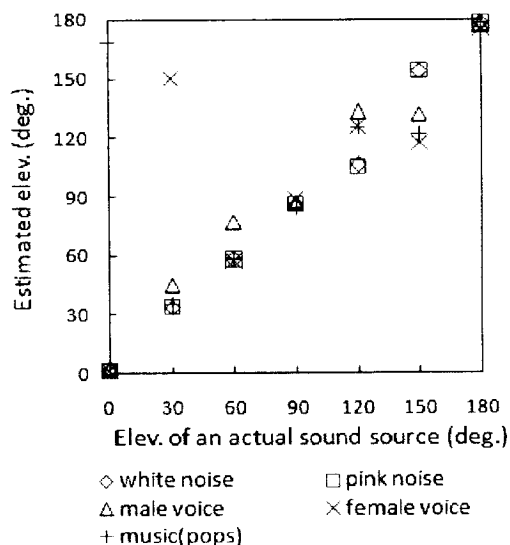


Fig.5 Estimated elevation for the sound source located in the upper median plane.

いることが関係していると思われる。これは、ヒトの音像定位の前後誤判断と類似の現象であり、興味深い結果である。また、同じ音源の異なる部分を用いると精度よく前方に推定できる場合もあり、前後判定の不安定さを示しているともいえる。さらに検討を進めたい。

参考文献

- [1] K. Iida *et al.*, Applied Acoustics, 68 (8), 835-850, 2007.
- [2] 例えば, H. Nakashima *et al.*, Acoust. Sci. Tech., 24 (4), 172-178, 2003.
- [3] 飯田他, 音講論 (秋), 295-296, 2000
- [4] K. Iida and M. Morimoto, J. Acoust. Soc. Am., 105(2) Pt.2, 1391 (1999)