

音声に対するピークホールド音源方向検出法の検討*

◎木皿 大介 金田 豊 (東京電機大・工)

1. はじめに

一般の室内で音源方向検出を行う場合、反射音の影響で検出性能が低下する。この問題に対して、これまで、信号の立ち上がり部分に着目した処理が検討されてきた[1-3]。その中で筆者らは、対数ピークホールド法を提案し、拍手音を用いて、様々な室条件でその有効性を検証してきた [4,5]。本報告では、拍手音に比べ、比較的音源方向検出が困難と考えられる音声を用い検討する。

2. 2つのマイクによる音源方向検出

音源は一方向と仮定して、2つのマイクロホンで信号を受音し、両信号間の時間差 τ_s を相互相関関数が最大値をとる時刻として検出する。そして、次式を用いて方向検出を行う。

$$\theta_s = \sin^{-1}(c \cdot \tau_s / d)$$

θ_s : 方向検出結果, c : 音速, d : マイク間距離

3. 対数ピークホールド法

前節で述べた方法は反射音が存在すると誤差を発生する場合がある。そこで、直接音の振幅にピークホールド処理を行うことで、後続する反射音列をマスクする。ただし、実際の音場では、反射音レベルが直接音レベルを超える場合がある。そこで振幅を対数値に変換して、その影響を軽減する[4]。

4. 処理手順[5]

- ①受音した信号波形の振幅に室内残響と同程度の指数減衰を持たせたピークホールド処理を行う。
- ②ピークホールド波形を対数値に変換する。
- ③信号の立ち上がり区間の検出：雑音がほぼ定常であると仮定し、短時間 T [ms] で A_{th} [dB] 以上の振幅変化があった時刻を信号の立ち上がりとみなした。(今回は音声を仮定しているので、 $T=20$ ms, $A_{th}=10$ dB とした)。

④次に、立ち上がり時刻の前後 50ms 程度の信号を切り出し、時間差分を取った後相互相関関数を計算して2節の方法により方向 θ_s を検出する。

5. 実験条件

本手法の反射音に対する有効性を確認するために、反射音の影響が大きいと考えられる、以下の音場条件で実験を行った。

- 配置：寸法が 5.0×9.0×2.4[m]、残響時間が 400ms の実験室においてマイクロホンの近くに反射板を図1のように設置した。マイクロホン間隔 d は 0.6m とした。
- 音源：男声3名による、「ばくおん」及び「おんきょう」の2種類の音声を用い、スピーカから再生した場合及び人間が発声した場合の2条件で、音源方向検出を行った。
- 音源位置：マイク-音源距離は 2,3,4m とし、音源方向は、各距離に対して1~4方向程度に設定した。

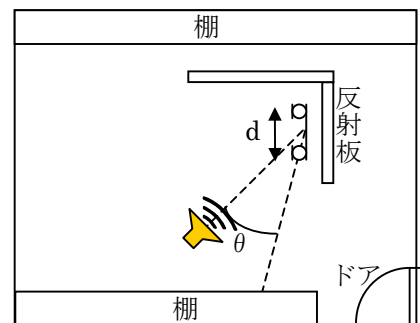


図1. 実験室でのマイクロホン対の配置図

6. 実験結果

2つの受音信号の相互相関関数のみから音源方向を検出する方法を従来法として比較を行った。

図2、図3は音源をスピーカとした場合の従来法と、提案法(対数ピークホールド法)の方向検出誤差を示している。検出誤差は5°以内、5°~10°の範囲、10°以上の3つの場合に分

*A study on DOA detection of a speech signal using peakholding, by KISARA, Daisuke and KANEDA, Yutaka (Tokyo Denki Univ.).

類し、それぞれの累計を、青、赤、黄の棒グラフとして表した。横軸はマイク-音源間距離を表す。

図 2、図 3 を比較すると、従来法（図 2）では誤差が 10° 以上ある部分が 50%であるのに対し、提案法（図 3）では誤差は 11%とほぼ正しく角度検出できており、提案手法の有効性が確認できる。

一方、図 4、図 5 は人間が発声した場合の結果を示している。この場合も従来法（図 4）は誤差が 10° 以上の部分が 56%をしめているのに対し、提案法（図 5）では 38%と誤差を軽減でき、有効性が確認できる。

これより、ピークホールド処理が音声の方向検出に対して改善効果があることが確認できた。しかし、 10° 以上の誤差も少なくはなく、更なる性能向上が望まれる。

ここで、図 2 と図 4 及び図 3 と図 5 を比較すると、スピーカから再生した結果より、人間が発声した場合の方が正しく角度検出ができていないのが確認できる。このように、同一音響条

件、同一音声でも音源が異なることで結果が異なっており、これは、スピーカと人間の指向特性の違いに原因があることが考えられる。

7. むすび

対数ピークホールド法は、反射音の影響が大きいと考えられる条件において、音声の到来方向をより少ない誤差で検出できることを示した。しかし、今後更に性能の向上が望ましい。また、音源の違いによる性能の違いについても、今後検討する。

参考文献

- [1]金田：音講論(秋), 547-548, (1991.10).
- [2]黄、他：信学論, J71-A, 10, 1780-1789, (1988.10).
- [3]小林、他：音講論(春), 535-536, (2001.3).
- [4]木皿、他：音講論(春), 611-612, (2005.3).
- [5]木皿、他：音講論(秋), 111, (2005.10).

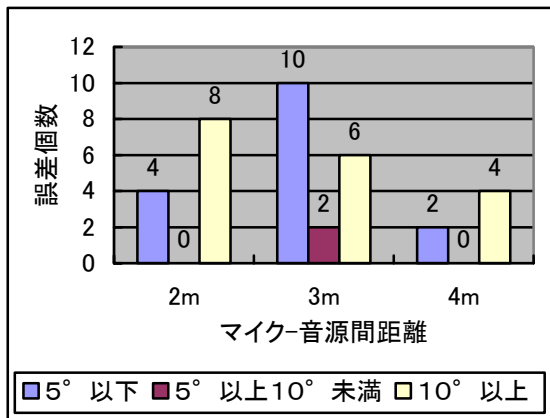


図 2. 従来法の誤差（音源：スピーカ）

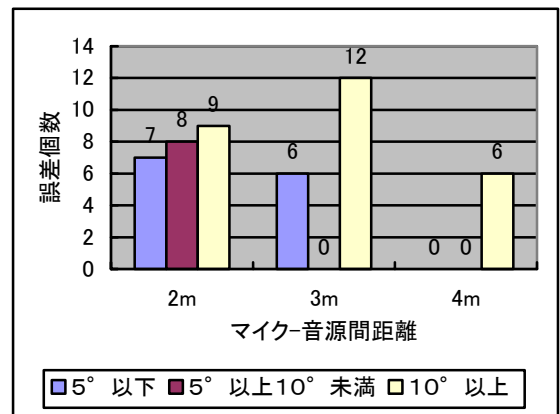


図 4. 従来法の誤差（人間発声）

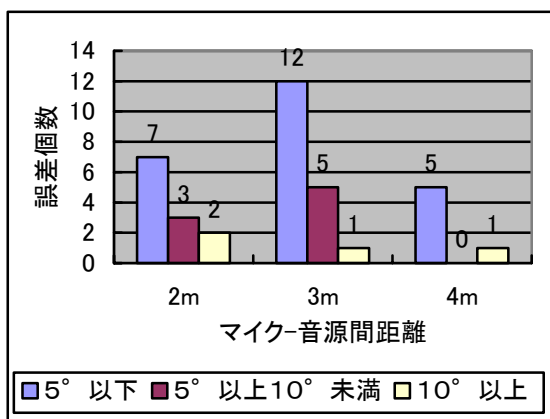


図 3. 提案法の誤差（音源：スピーカ）

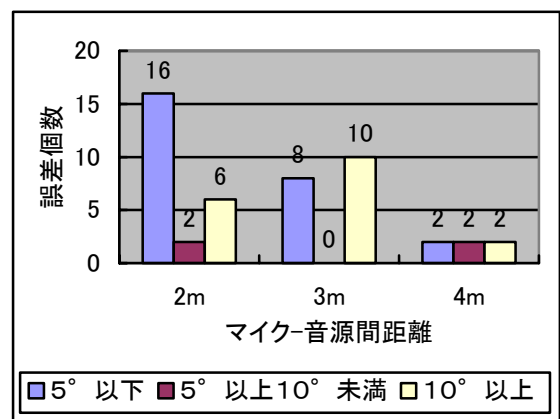


図 5. 提案法の誤差（人間発声）