

音源方向推定に及ぼす室内反射音の影響分析*

◎上杉信敏, 金田豊 (東京電機大・工)

1 はじめに

室内で音源方向推定技術を利用する場合には、反射音の影響を無視することはできない。本報告では最も基本的な 2ch マイクロホンアレーを対象とし、鏡像法によるシミュレーションを行うことで室内反射音の影響評価を行ったので報告する。

2 音源方向推定法

2 つのマイクロホンを用いて音源方向を推定するための基本的な手法は、受信信号 $x_1(t)$ と $x_2(t)$ 間の遅延時間を、次式で定義される相互相関関数(CC: Cross Correlation function) $\phi_{12}(\tau)$ の最大値から推定する方法である。

$$\phi_{12}(\tau) = E[x_1(t)x_2(t+\tau)] \quad (1)$$

ここで $x_1(t)$ と $x_2(t)$ のクロススペクトル $\Phi_{12}(\omega)$ を用いると、式(1)の $\phi_{12}(\tau)$ は次式で表される。

$$\phi_{12}(\tau) = \int \Phi_{12}(\omega) e^{j\omega\tau} d\omega \quad (2)$$

一方、周囲雑音や室内反射音などの外乱下での性能を確保するため $\Phi_{12}(\omega)$ に様々な周波数重み $\Psi(\omega)$ を付ける事が提案されてきた[1]。これを一般的に表したものが一般化相互相関関数 $R_{GCC}(\tau)$ で、次式で表される。

$$R_{GCC}(\tau) = \int \Psi(\omega) \Phi_{12}(\omega) e^{j\omega\tau} d\omega \quad (3)$$

代表的な重み関数 $\Psi(\omega)$ を Table 1 に示す。

3 反射音の影響評価

3.1 評価条件

反射音が方向推定に及ぼす影響を評価するために鏡像法[2]に基づく計算機シミュレーションを行った。想定する部屋のサイズを $9.0 \times 5.0 \times 2.4$ m とし、2つのマイクロホン間隔を 0.6 m とした。様々な条件の反射音を考慮するため、Fig. 1 に示す音源距離 R を 3 m、音源方向 θ_s を 60° に固定しながら、アレーの配置(アレーの中心位置 M_c およびアレーの向き θ_M)をランダムに 5000 パターン変化させた。残響時間が 0~0.4 s の範囲で変化するように壁面等の反射率を変化させた。その他のシミュレーション条件は Table 2 に示した。

Table 1 代表的な周波数重み $\Psi(\omega)$

推定法	周波数重み $\Psi(\omega)$
CC	1 : 通常の相互相関関数
SCOT	$\frac{1}{\sqrt{ \Phi_{11}(\omega)\Phi_{22}(\omega) }}$: 振幅をコヒーレンス関数に比例させる
PHAT	$\frac{1}{ \Phi_{12}(\omega) }$: 信号の白色化
ML	$\frac{1}{ \Phi_{12}(\omega) } \cdot \frac{ \gamma_{12}(\omega) ^2}{1- \gamma_{12}(\omega) ^2}$: $\gamma_{12}(\omega)$ コヒーレンス関数

Table 2 シミュレーション条件

サンプリング周波数	12000 [Hz]
スペクトルを求めるための周波数分析フレーム長	1024 [point]
フレームシフト長	256 [point]
窓関数	ハミング窓
目的音	男性音声 (250 [ms])

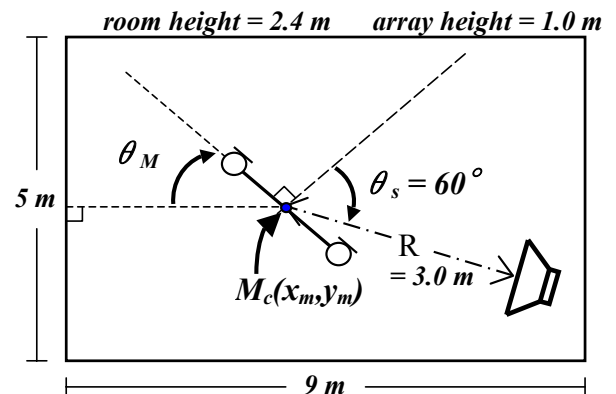


Fig. 1 マイクロホンアレー配置

3.2 各推定法での反射音影響の比較

許容誤差を $\pm 5^\circ$ として求めた各手法の方向推定正解率を Fig. 2 に示す。Fig. 2 が示すように、CC 法は残響時間が 0.05 s、SCOT、PHAT、ML 法は 0.25 s 程度まで良好に推定できた。しかし、残響時間が長くなるにつれ、各推定法の性能は劣化した。特に CC 法の劣化は著しく残響時間 0.4 s では推定正解率 5% となった。それに対し、残響時間 0.4 s において SCOT、PHAT 法は 80%、ML 法は 75% であった。このことより、反射音に対しては、SCOT、PHAT 法 > ML 法 >> CC 法の順に性能が良いことがわかった。

* Effect of Room Echoes on DOA estimation, by UESUGI Nobutoshi and KANEDA Yutaka (Tokyo Denki Univ.).

3.3 CC 法に及ぼす反射音影響

CC 法が他手法に比べ性能が劣る理由を検討した。Fig. 3 は音源-各マイク間のインパルス応答の相互相関関数の例を表す。音源方向に対応する直接音同士の相関ピーク以外にも、反射音の影響で相関の高いピークが複数存在した。

CC 法ではこの相互相関関数に音声(音源信号)の自己相関関数を畳み込んだもの(Fig. 4)が得られる。Fig. 3 における時間遅れが-5 サンプル付近のように相関ピークの密度が高い(図では複数のピークが融合して、その幅が広がっている)部分の相関値が、Fig. 4 では 18.5 サンプル付近の正解ピークより大きくなり、誤推定を発生させる。

これに対して、SCOT, PHAT 法は音声を白色化して相互相関関数をインパルス応答のそれに近いものとするため、Fig. 3 と類似した結果が得られる。このことで、個々の相関ピークの融合は減少し、直接音同士の相関ピークが最大となる。

3.4 壁面からの距離の影響

アレーを壁際に置いた場合には推定性能の劣化が予測される。そこで、マイクロホンアレー位置と誤推定率の関係を検討した。

残響時間 0.4 s で PHAT 法の場合の誤推定率を、アレー中心から壁面までの距離毎に分類し、Fig. 5 に示す。Fig. 5 が示すように、誤推定はアレーが壁面に近づくにつれ大きくなる。しかし、壁面から 1.0 m 以上ではほぼ一定の誤推定率を示す。この原因は、アレーから壁面までの距離が床・天井までの距離約 1.0 m を越えると、床・天井からの反射の影響が主となり、この影響は壁面からの距離に依存しないためだと考えられる。

4 まとめ

本報告では、鏡像法による残響シミュレーションを用いて、2 マイク信号間の時間差検出に基づく音源方向推定に及ぼす反射音の影響を検討した。その結果、SCOT, PHAT 法などは音声信号を白色化して相互相関のピークを鋭くするため、反射音の影響で発生するピークとの融合を回避することができ、通常の相互相関法と比較して誤推定率が小さいことがわかった。

参考文献

- [1] C. H. Knapp et al, IEEE Trans. on Acoust., Speech and Signal Proc., ASSP-24, 4, 320-327, Aug. 1976.
- [2] J. Borish, J. Acoust. Soc. Am., 75, 1827-1836, Jun. 1984.

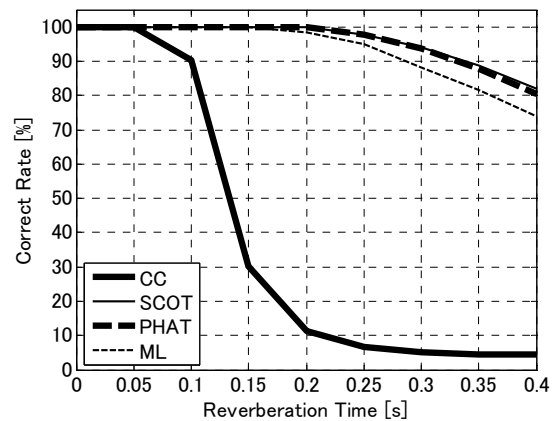


Fig. 2 各推定法での反射音影響の比較

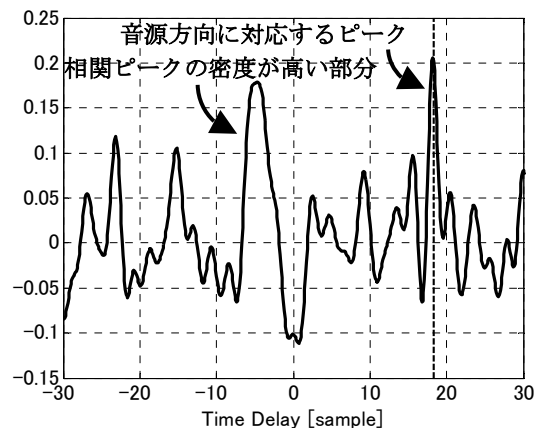


Fig. 3 インパルス応答の相互相関関数

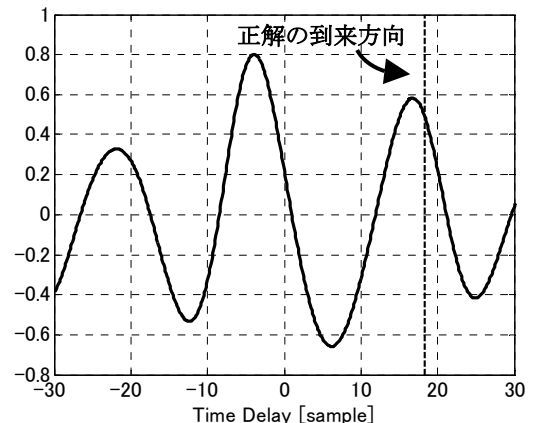


Fig. 4 音声に対する相互相関関数

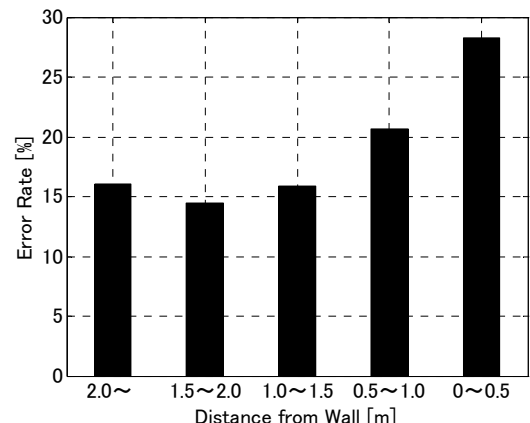


Fig. 5 壁面からの距離に対する誤推定率