

MUSIC 音源方向推定法におよぼす反射音の影響と SBPHP 処理の検討*

☆湯浅 一輝, 金田 豊 (東京電機大・工)

1 はじめに

音源方向推定法には、多チャンネル信号を用いる MUSIC 法[1]をはじめ、様々な手法が存在する。しかし、室内環境での音源方向推定では、反射音が推定性能の劣化を引き起こす。加藤らは、反射音への対策法として、SBPHP (Sub Band Peak Hold - Phase) 処理を MUSIC 法に適用した SBPHP-MUSIC 法を提案し、反射音耐性向上の確認を行った[2]。

本報告では、シミュレーションにて PH 処理の有効性を確認し、実録音声にて MUSIC 法と SBPHP-MUSIC 法の推定性能を比較した。

2 MUSIC 法による音源方向推定

方向推定を行う場合、各マイクの受信信号には音源方向 θ に依存した時間差 τ が生じる。第 1 チャンネルを基準として、 θ 方向から到来する音の多チャンネル受信信号は周波数領域で次式のように表される。

$$\begin{aligned} \mathbf{X}(\omega, \theta) &= X_1(\omega) \left[1, e^{-j\omega\tau_2(\theta)}, e^{-j\omega\tau_3(\theta)}, \dots \right]^H \\ &= X_1(\omega) \mathbf{d}(\omega, \theta) \end{aligned} \quad (1)$$

ここで、 $X_1(\omega)$ は基準とした信号スペクトル、 ω は角周波数、 $\mathbf{d}(\omega, \theta)$ は方向制御ベクトル (steering vector) を表している。

MUSIC 法では、受信信号から算出される雑音部分空間の相関行列 $\mathbf{R}_n(\omega)$ と $\mathbf{d}(\omega, \theta)$ を用いて、次式によって方向スペクトル $P_{MUSIC}(\omega, \theta)$ を求める。

$$P_{MUSIC}(\omega, \theta) = \frac{1}{\mathbf{d}^H(\omega, \theta) \mathbf{R}_n(\omega) \mathbf{d}(\omega, \theta)} \quad (2)$$

3 サブバンドピークホールド (SBPH) 処理

SBPH 処理は、受信信号を短時間フーリエ変換するサブバンド処理と、周波数帯域毎に先行する直接音を利用し、後続の反射音をマスクするピークホールド処理で成り立つ。

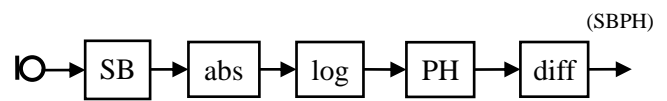


図 1 SBPH 処理のブロック図

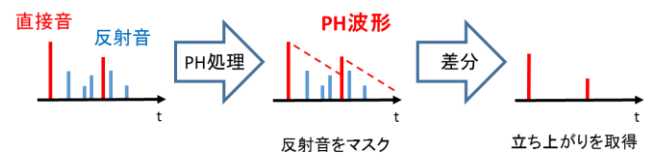


図 2 PH 処理のモデル図

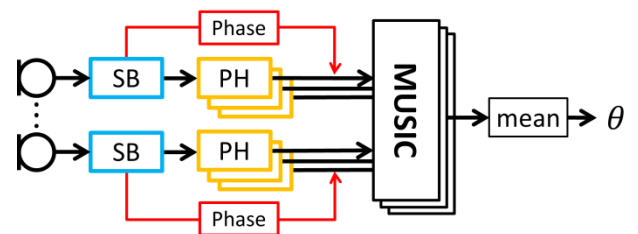


図 3 SBPHP-MUSIC 法のブロック図

SBPH 処理の流れを説明する(図 1)。最初に短時間フーリエ変換によって信号をサブバンド化:SB する。次に各サブバンド信号の振幅成分:abs の対数値:log を取り出し、ピークホールド処理:PH(図 2)を行う。最後に時間差分:diff をとる (以下、この abs から diff までの 4 つの処理をまとめて PH 処理と呼ぶ)。その結果、周波数帯域毎に直接音成分のみが抽出される。

4 SBPHP-MUSIC 法

SBPHP-MUSIC 法では、図 3 に示すように PH 処理前の信号の持つ位相情報をあらかじめ保持しておく、PH 処理後の直接音成分のみの信号に位相を付加することで各信号の位相差は回復する。この位相を付与する処理を phase 処理という。

Phase 処理を行うことで PH 処理により失われた位相成分を取り戻し、MUSIC 処理への適用が可能となる。結果、直接音だけの位相差で計算を行うことができる。

* A study of effects of reflective sounds and SBPHP processing on MUSIC sound source direction estimation, by Kazuki YUASA, and Yutaka KANEDA (Tokyo Denki University).

5 PH 処理の有効性検証

シミュレーションによって PH 処理の反射音抑圧性能を検証した。条件は音声が入射が 45° 方向から到来し、その反射音が -60° 付近から複数到来するものである(図 4, 表 1)

MUSIC 法および SBPHP-MUSIC 法による推定結果を図 5(a)(b)に示す。図 5(a)の MUSIC 法の推定結果では、複数の反射音が融合して -60° 付近に誤った推定ピークが発生している。一方、PH 処理を適用した図 5(b)は、複数の反射音が抑圧されて直接音方向のピークのみが表れている。

このように PH 処理は反射音抑圧効果があり、推定性能が向上することを確認できた。

6 実音場での評価実験

実録音源にて MUSIC 法と SBPHP-MUSIC 法の誤推定比較実験を行った。実験は以下の条件で行われた。

条件①(近距離)：比較的反射音が少ない表 2 で表される条件。

条件②(近距離反射強)：表 2 の条件で、アレー位置を 1m から 0.28m に変化させ、壁からくる初期反射音成分を強くした条件。

条件③(遠距離)：表 2 の条件で発話距離を 4, 5m にする。これにより直接音成分が小さくなり、相対的に反射音が大きくなる条件。

結果を表 3 に示す。近距離条件では両推定法とも誤推定は無かった。しかし、反射音の影響が大きい条件②と条件③では、MUSIC 法では多くの誤推定が発生したのに対し、PH 処理を適用した SBPHP-MUSIC 法では誤推定数を大幅に減少することを確認できた。これはシミュレーションによる検証と同様に反射音成分を抑圧できている結果と思われる。

7 まとめ

本報告では、シミュレーションを使用し、PH 処理が反射音の抑圧に有効であることを確認した。また、複数の実音場条件において MUSIC 法と SBPHP-MUSIC 法との誤推定比較を行った。結果、PH 処理による誤推定の減少を確認した。

参考文献

- [1] 浅野、“音のアレイ信号処理”，日本音響学会、115-124, 2011.
- [2] 加藤、金田、音講論(秋), 731-732, 2012.

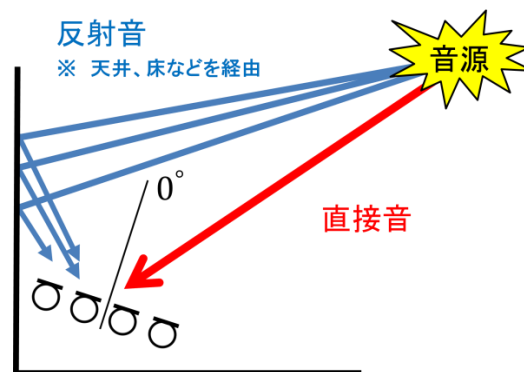


図 4 シミュレーションのイメージ図

表 1 シミュレーションにおける反射音の条件

反射音方向 [$^\circ$]	直接音との振幅比	遅延時間 [ms]
-60	0.9	3
-55	0.8	3.3
-65	0.7	3.6

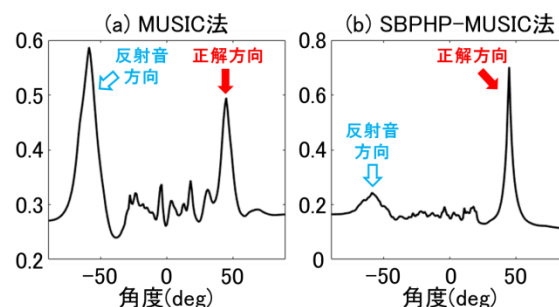


図 5 方向推定シミュレーションの結果

表 2 実験条件(実験条件①:近距離)

実験室	6.3×9.1×2.8(WDH) m
残響時間	1.3s (500Hz)
マイクロホンアレー	4ch, 4cm 間隔
アレー位置	部屋の隅から 1m
サンプリング周波数	48000Hz
発話距離	1, 2, 3m
発話角度	0, ± 15 , ± 30 , $\pm 45^\circ$
許容誤差	$\pm 5^\circ$
使用音声データ	男性 1 名, 10 単語

表 3 SBPHP 処理有無での誤推定数

	音声数 (個)	MUSIC	SBPHP MUSIC
①:近距離	210	0	0
②:近距離反射強	210	16	3
③:遠距離	140	42	7

※各手法、各条件で最適な帯域制限を行った。
(例：条件① MUSIC 3~12.5[kHz]
SBPHP-MUSIC 2.5~12.5[kHz])