

サブバンドピークホールド処理の適用による MUSIC法の反射音耐性向上について*

◎鈴木敬, 金田豊 (東京電機大・工)

1 はじめに

一般の室内で音源方向推定を行う際, 反射音の影響により推定精度が低下する。この対策として, 我々は, サブバンドピークホールド (SBPH) 処理を提案し, 2 マイクに基づく相互相関法に適用して, その有効性を確認した [1]。本報告では, 複数マイクを用いた MUSIC 法に SBPH 処理を適用することで, 反射音耐性の向上が確認されたので報告する。

2 MUSIC 法 [2] による音源方向推定

M 個のマイクロホンで受音した信号を中心周波数 ω で狭帯域化し, マイクロホン i, j 間でのクロススペクトル $\Phi_{ij}(\omega)$ ($i, j=1, 2, \dots, M$)

$$\Phi_{ij}(\omega) = E[X_i^*(\omega)X_j(\omega)] \quad (1)$$

を求める。ただし, *は複素共役, $E[\cdot]$ は期待値を表す。

そして, $\Phi_{ij}(\omega)$ を ij 要素とした空間相関行列を $\mathbf{R}(\omega)$ と表す。この空間相関行列 $\mathbf{R}(\omega)$ の固有空間構造 (信号部分空間と雑音部分空間の直交性) を利用して, 周波数 ω での音源方向スペクトル (音源方向推定結果) $P(\omega, \theta)$ を次式のように求める。

$$P(\omega, \theta) = 1/\{\mathbf{d}^H(\omega, \theta)\mathbf{R}_N(\omega)\mathbf{d}(\omega, \theta)\} \quad (2)$$

$$\mathbf{R}_N(\omega) = \sum_{q=L+1}^M \mathbf{v}_q(\omega)\mathbf{v}_q^H(\omega) \quad (3)$$

ただし, H は共役転置, $\mathbf{v}_q(\omega)$, $q=L+1, \dots, M$ (L : 音源数) は雑音部分空間を張る固有ベクトルを表す。また, $\mathbf{d}(\omega, \theta)$ は次式で定義する方向制御ベクトル (steering vector) を表す。

$$\mathbf{d}(\omega, \theta) = [1, e^{-j\omega\tau_2(\theta)}, \dots, e^{-j\omega\tau_M(\theta)}]^T \quad (4)$$

ただし, T は転置を表し, $\tau_i(\theta)$ は θ 方向から到来する音が i 番目のマイクに生じる遅延時間 (相対値) を表す。

最後に, 式(5)に示すように, $P(\omega, \theta)$ を指定

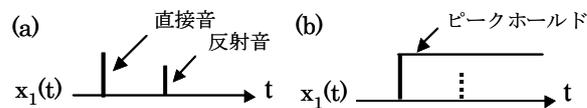


図1 ピークホールド処理の説明図

した周波数範囲 $\omega_1 \sim \omega_2$ で加算して音源方向スペクトル $P(\theta)$ を得る。

$$P(\theta) = \sum_{\omega=\omega_1}^{\omega_2} P(\omega, \theta) \quad (5)$$

3 サブバンドピークホールド (SBPH) 処理 [1] [3]

図1 (a)に, パルス音に単一反射音が付加された受音波形 $x_1(t)$ の例を示し, (b)に受音波形にピークホールドをかけた波形を示す。このように, ピークホールド処理は直接音の大きさを保持することで, 後続する低振幅の反射音をマスクする。

また, 本処理では高振幅の反射音を想定し, ピークホールド後に対数をとって高振幅反射音の影響を軽減する (対数操作)。その後, 時間差分により, ピークホールド信号の持つ低域成分を抑圧する。そして, 帯域毎に到来時刻が異なる音声に対応させるために, 信号をサブバンド信号に分割した後, 以上のピークホールド処理を行う (SBPH: Sub-Band Peak Hold)。

4 提案手法 (SBPH - MUSIC)

SBPH処理をMUSIC法に適用した処理系のブロック図を図2に示す。まず, 受音信号を短時間フーリエ変換 (STFT) し, サブバンド信号の振幅成分の時系列 $|X_i(k, t)|$ (i : マイク番号, k : サブバンド番号, t : 時間) を出力する。その後, ピークホールド処理 (PH), 対数操作 (\log), 時間差分 (Diff) を行って, 反射音の影響を軽減した信号を $|X_i(k, t)|_p$ と表す。この時, 各チャンネルの k 個のサブバンド信号 $|X_i(k, t)|_p$ は, 各々狭帯域化された時間波

* Improvement of the robustness of MUSIC method to reflected sound by using Sub-band Peak Hold processing, by SUZUKI, Takashi and KANEDA, Yutaka (Tokyo Denki University).

形となっている。

次に、第 k 番目のサブバンドにおいて (サブバンド番号を固定値にして), 2つの時間波形 $|X_i(k,t)|_p$ と $|X_j(k,t)|_p$ とのチャンネル間のクロススペクトル $\Phi_{ij}(\omega',k)$ を次式から計算する。

$$\Phi_{ij}(\omega',k) = E[F[|X_i(k,t)|_p]^* \cdot F[|X_j(k,t)|_p]] \quad (6)$$

ただし、 $F[\cdot]$ はフーリエ変換、 ω' は各 $|X_i(k,t)|_p$ 信号の周波数を表す。

これを全てのサブバンドに対して行い、 k に関する総和を求めたものを $\Phi_{ij}(\omega')$ と表す。

$$\Phi_{ij}(\omega') = \sum_k \Phi_{ij}(\omega',k) \quad (7)$$

次に、この $\Phi_{ij}(\omega')$ を ij 要素とした周波数 ω' の空間相関行列を $\mathbf{R}(\omega')$ と表す。以下、第 2 章の式(2)-(5)と同様に $\mathbf{d}(\omega',\theta)$ と $\mathbf{R}(\omega')$ より求めた $\mathbf{R}_M(\omega')$ を用いて、音源方向スペクトル $P(\theta)$ を計算する。

5 実環境下における評価実験

最初に表 1 に示す実験条件で、アレーを部屋の中心付近に配置して第 1 の実験を行った。その後、初期反射音の影響が大きい条件として、部屋の角に近接するようにアレーを配置して (図 3) 第 2 の実験を行った。

音源は人間 (男性 1 名) が 30 単語を、各単語・角度 (3 方向) 毎に 3 回ずつ発声し、合計 270 条件で測定した (第 1 の実験では各 1 回で合計 90 条件)。以上の条件で、通常の MUSIC 法と SBPH-MUSIC 法との比較を行った。

実験結果の正解率 (許容誤差 10deg.) を図 4 に示す。(a) が第 1 の実験、(b) が第 2 実験の正解率である。図より、初期反射音の影響の少ない第 1 の実験では、従来法(MUSIC)と提案法(SBPH - MUSIC)の両方ともに 100% の正解率が得られた。一方、初期反射音の影響が大きい第 2 の実験では、従来法の正解率が 85% まで劣化した。これに対して、提案法は約 99% と高い正解率を維持しており (ただし、許容誤差 5deg. の正解率では 96%), 反射音耐性の向上が確認できた。

6 むすび

本報告では、反射音耐性のあるサブバンドピークホールド処理を MUSIC 法に適用した。その結果、初期反射音の影響が大きい室内条件であっても通常の MUSIC 法に比べ、高い音源方向推定性能を維持できることを示した。

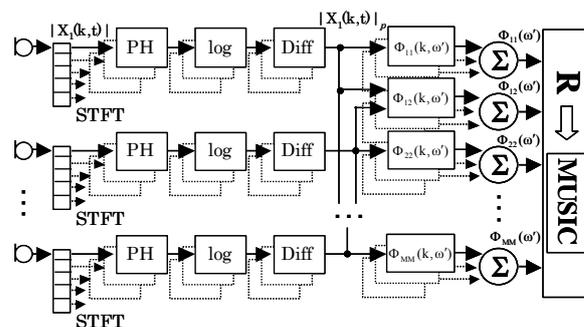


図 2 提案手法のブロック図

表 1 実験条件

部屋の寸法	5.0[W]×6.0[D]×2.5 [H] [m]
残響時間	0.38 s
SN 比	25 ~ 30 dB
アレーの高さ	1.2 m
マイクロホン数	8
アレーの形状	円形
アレーの直径 d	0.3 m
音源距離 r	3.0 m
音源方向 θ_s	0, 30, 45 deg.
標本化周波数	32000 Hz
SBPH 処理の分析フレーム長	$T = 32$ [sample]
分析シフト長	$T/8$ [sample]
MUSIC 法の使用周波数 ω	500 ~ 3000 Hz
SBPH-MUSIC 法の使用周波数 ω'	500 ~ 2000 Hz

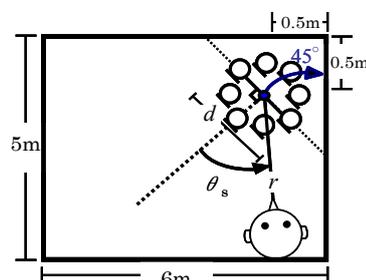
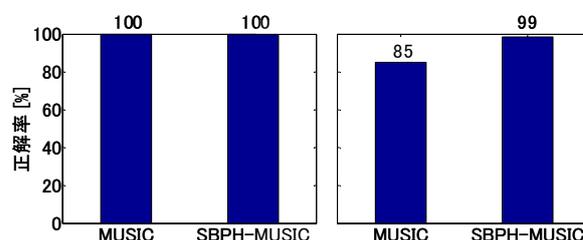


図 3 本実験の配置図



(a) 第 1 の実験 (初期反射音小)
(b) 第 2 の実験 (初期反射音大)

図 4 音源方向推定実験の正解率

参考文献

- [1] 鈴木, 金田, 音講論(秋), 751-752, 2007.
- [2] 大賀, 山崎, 金田, "音響システムとデジタル処理," 電子情報通信学会, 1995.
- [3] 木皿, 金田, 音講論(春), 631-632, 2006.