

サブバンドピークホールド処理を用いた 多チャンネル音源方向推定法の評価*

○霜田皓之, 金田豊 (東京電機大・工)

1 はじめに

音源方向推定を行う際の誤差要因として、反射音と雑音がある。我々は反射音対策として、サブバンドピークホールド(SBPH)処理を2チャンネル音源方向推定法である相互相関(CC)法に適用したSBPH-CC法の検討を進め、大幅な反射音耐性向上を実現した[1]。しかし、雑音耐性は十分な性能を満たしていないことが問題として挙げられていた。

本稿ではSBPH処理を多チャンネル手法の遅延和ビームフォーミング(Delay and Sum Beamforming 以下DS法)法[2]に適用することで、雑音耐性向上が確認されたので報告する。

2 サブバンドピークホールド処理(SBPH)

図1に、パルス音に単一反射音が付加された受信波形と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。

次に、サブバンドピークホールド(SBPH: Sub-Band Peak Hold)処理の流れを図2に従って説明する。まず、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に分割する。その後振幅(abs)のピークホールド(PH)を行った後、対数値の時間差分(Diff)をとって音の立ち上がりを得る。

3 SBPH-DS法(提案法)

DS法は各マイクロホンで受信した信号に遅延を付加し、方向 θ の音を同相化する手法である。このDS法にSBPH処理を適用したSBPH-DS法の処理の流れを図2に示した。SBPH処理を通した信号を $X_i(k, t)$ (i :マイク番号 k :サブバンド番号)とする。次に、 k 番目のサブバンドにおいて各チャンネル間のクロススペクトル Φ_{ij} を次式により求める。

$$\Phi_{ij}(k, \omega) = E[F[X_i(k, t)]^* \cdot F[X_j(k, t)]] \quad (1)$$

ただし、 $*$ は複素共役、 $E[\cdot]$ は期待値、 $F[\cdot]$ はフーリエ変換、 ω は各 $X_i(k, t)$ 信号の周波

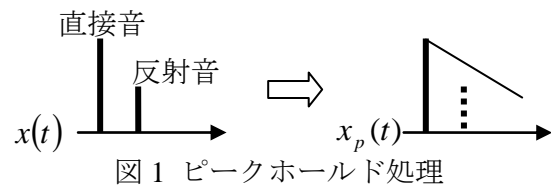


図1 ピークホールド処理

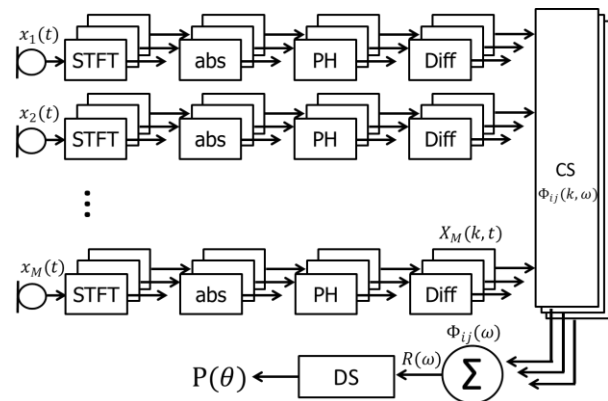


図2 SBPH-DS法のブロック図

数を表す。これを全ての k に対して行い、総和を求めたものを $\Phi_{ij}(\omega)$ とする。この $\Phi_{ij}(\omega)$ を ij 要素とした周波数 ω の空間相関行列を $R(\omega)$ とし、 $\omega_1 \sim \omega_2$ で帯域制限して次式よりSBPH-DS法の音源方向スペクトル $P(\theta)$ を計算する。

$$P(\theta) = \sum_{\omega=\omega_1}^{\omega_2} d^H(\omega, \theta) R(\omega) d(\omega, \theta) \quad (2)$$

ただし、 H は共役転置で、方向制御ベクトル $d(\omega, \theta)$ は以下の式で与えられる。

$$d(\omega, \theta) = \begin{bmatrix} e^{-j\omega\tau_1(\theta)} \\ \vdots \\ e^{-j\omega\tau_M(\theta)} \end{bmatrix} \quad (3)$$

ここで、 $\tau_i(\theta)$ は θ 方向から音が到来した時の、 i 番目のマイクに生じる遅延時間である。

4 実験条件

実験は図3の環境において人間の発話の録音を行い、表1の実験条件にて、SBPH-CC、DS、SBPH-DS法の雑音耐性と反射音耐性の評価

* Evaluation of multi-channel sound source direction estimation using Sub-Band Peak Hold processing, by SHIMODA, Hiroyuki and KANEDA, Yutaka (Tokyo Denki University).

表 1 実験条件

残響時間	0.38s
音源信号	ランダム日本語単語音声 (20 種)
音源方向	0° ~60° (30° 間隔)
マイク数	SBPH-CC : 2ch DS, SBPH-DS : 4ch
アレー形状	等間隔円形
アレー直径	0.3m
音源距離	0.5m~3.0m (0.5m 間隔)
SN 比	0dB~30dB (5dB 間隔)

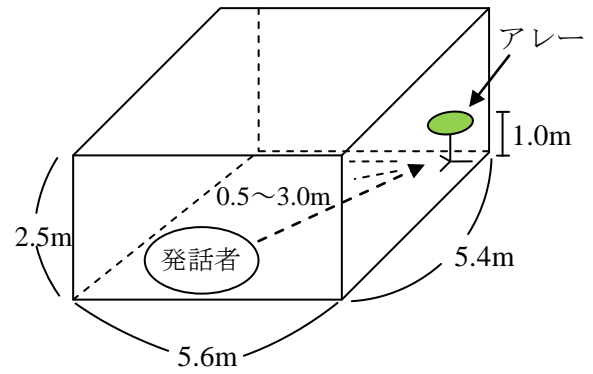


図 3 実験環境

を行った。この時、雑音耐性を評価するために、録音された多 ch 音声に平均的な室内雑音スペクトルを持つ拡散性 Hoth 雑音を付加し、0 ~30dB の間で SN 比を変化させた。また、音源距離が広がると直接音のエネルギーが低くなり反射音の影響が大きくなる現象を利用し、音源距離を 0.5~3.0m の間で変化させて反射音耐性を評価した。

5 評価実験

SBPH-CC, DS, SBPH-DS 法の実験結果をそれぞれ図 4 の(a), (b), (c)に示した。図は横軸が SN 比、縦軸が音源距離、色彩で正答率を表示した。また、比較のため、SBPH-DS 法の結果上に SBPH-CC 法の正答率のライン(90%)を破線にて表示した。

図4(c)において、破線で示した SBPH-CC 法の 90%正答率のラインよりも SBPH-DS 法の 90%ラインは右側に移動している。これは雑音の影響が強くなっても SBPH-CC 法よりも SBPH-DS 法が高い正答率を維持できることを意味しており、雑音耐性が向上していると言える。

次に実験結果(b), (c)を比較する。DS 法の正答率のラインが SBPH-DS 法では大きく持ち上がっている。そのため、反射音耐性が大幅に向上していることが言える。

これらの結果より、SBPH 処理を DS 法に適用することで SBPH-CC 法と同程度の反射音耐性を保ったまま、雑音耐性が向上することが確認できた。

6 むすび

SBPH 処理を DS 法に適用する SBPH-DS 法を提案した。その結果、従来法である SBPH-CC 法と同程度の反射音耐性を保ったまま、雑音耐性が向上することを確認した

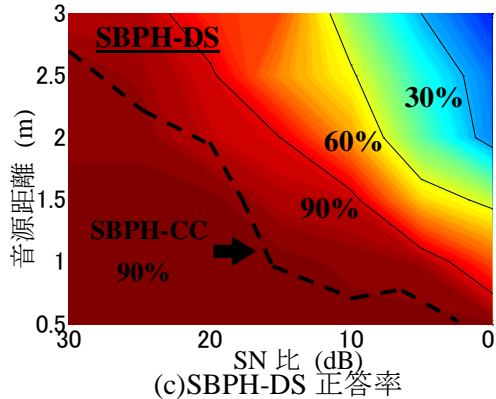
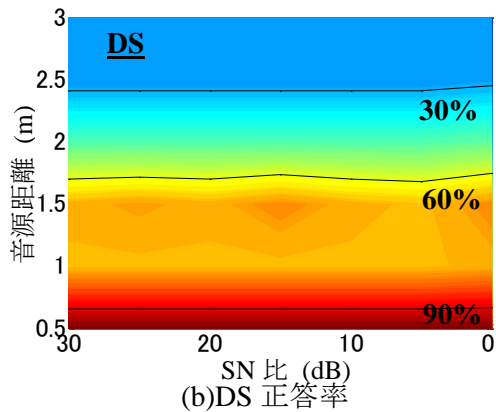
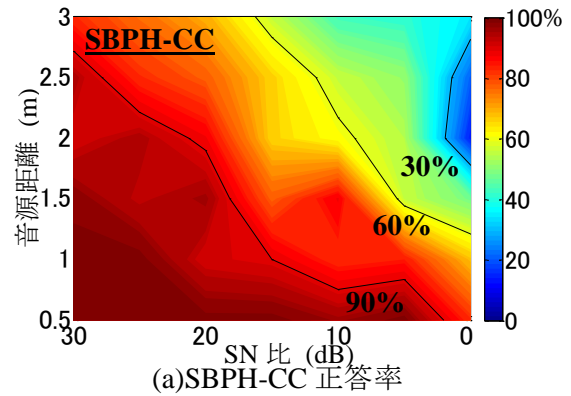


図 4 実験結果

参考文献

- [1] 鈴木ほか, 日本音響学会誌, 65, 10 (2009).
- [2] 大賀ほか, “音響システムとデジタル信号処理,” 電子情報通信学会 (1995).