

SBPHP-MUSIC 音源方向推定法における サブバンド化パラメータの検討

Study of the sub-band parameters in Sub-band Peak Hold Phase MUSIC DOA Estimation

朝比奈駿 金田豊
Shun Asahina Yutaka Kaneda
東京電機大学大学院 工学研究科
Graduate School of Engineering, Tokyo Denki University

1. はじめに

室内で音源方向推定を行う際、反射音の影響により推定性能が劣化する。筆者らは先に、反射音耐性に優れるSBPHP(Sub-band Peak Hold Phase)処理を従来の代表的手法であるMUSIC法と組み合わせたSBPHP-MUSIC法を提案し、MUSIC法より精度向上を確認した[1]。本報告では、SBPHP処理におけるサブバンド化パラメータに関する検討結果を述べる。

2. MUSIC 法

θ 方向から到来する音を多チャンネルで受信することを考える。第 1 チャンネルを基準として、各チャンネルの受信信号の遅延時間を $\tau_i(\theta)(i = 2,3, \dots)$ と表すと、方向制御ベクトルが次式で定義される。

$$\mathbf{d}(\omega, \theta) = [1, e^{j\omega\tau_2(\theta)}, e^{j\omega\tau_3(\theta)}, \dots]^H \quad (1)$$

この時、MUSIC 法の方角スペクトル $P_{MUSIC}(\omega, \theta)$ は、受信信号から算出した相関行列の雑音部分空間である $\mathbf{R}_n(\omega)$ と $\mathbf{d}(\omega, \theta)$ を用いて、次式により求められる。

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

3. SBPHP-MUSIC 法

PH 処理は、先行する直接音の振幅値を残響相当の減衰を持たせて保持することで、後続する反射音をマスクする。その後、対数を取り、時間差分をとることで、音源方向情報を含んだ直接音だけが取り出される(図 1)。

処理の流れを図 2 に示す。まず、帯域毎に立ち上がり時刻が異なる音声信号から効率的に直接音成分を取り出すために、マイクロホンで受信した音声信号を、短時間フーリエ変換(STFT)によりサブバンド化する。次に絶対値(abs)を取った後、PH 処理、対数操作(log)、時間差分(diff)によって、帯域毎に音声信号の直接音成分を得る。次に位相情報を付加し、式(2)の MUSIC 方向スペクトルを計算した後、複数の周波数に対して平均化を行い、音源方向を推定する。

4. パラメータの検討

SBPHP-MUSIC 法のサブバンド処理時のフレーム長とシフト長と推定誤差数の関係を検討した。実験は容積が約 $100m^3$ 、残響時間 0.4s の部屋で、マイク 4 個を 0.05m 間隔で直線上に配置した。発話者 1 名、音声は日本語 12 単語、9 方向($0^\circ, \pm 15^\circ, \pm 45^\circ, \pm 60^\circ$)、発話距離は 1, 2, 3m で総発話数は 324、許容誤差 $\pm 5^\circ$ 。サンプリング周波数は 48kHz とした。

実験結果を表 1 に示す。(フレーム長, シフト長)=(1024,256)が誤推定数, 処理時間から最も良好な結果が得られた。また、同条件で MUSIC 法と比較した(表 2)。結果から誤推定数が 1/2 に減少しており、SBPHP 処理の有効性が保たれていることを確認した。

5. まとめ

本報告では、反射音を抑制する SBPHP-MUSIC 法におけるサブバンド化パラメータの検討を行った。今回の実験条件ではフレーム長 1024, シフト長 256 で最も性能が良く、MUSIC 法と比較した場合、誤推定数は 1/2 に減少した。

謝辞 本研究の一部はJSPS 科研費 15H02728 の助成を受けたものです。

参考文献

- [1] 加藤 他, 音講論集(秋), 731-732, (2012).

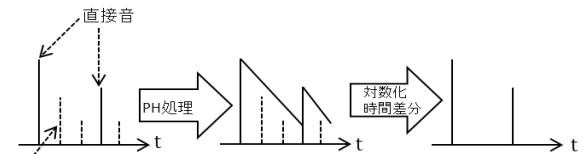


図 1 ピークホールド処理の原理図
位相付加

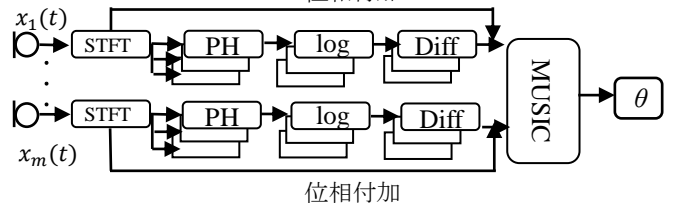


図 2 SBPHP-MUSIC 法のブロック図

表 1 フレーム長, シフト長と誤推定数(全 324 データ)

Frame \ Shift	256	512	1024	2048
1/8frame	13(1.04)	15(1.10)	4(1.37)	4(1.91)
1/4frame	13(0.61)	13(0.72)	4(1.01)	5(1.55)
1/2frame	14(0.41)	13(0.73)	5(0.79)	5(1.36)

表 2 推定性能比較

	誤推定数
SBPHP-MUSIC 法	4(1.01)
MUSIC 法	9(0.98)

※表内の()内は 1 発声
当たりの処理時間[s]