

線形予測に基づく初期反射音抑圧を用いた音源方向推定について*

◎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、まず、反射音が推定法に及ぼす影響を分析する。そして、その結果に基づいて初期反射音の抑圧を行う新しい方向推定手法を提案し、性能評価を行った結果を報告する。

2 反射音の影響分析

反射音が音源方向推定に及ぼす影響を分析するために鏡像法[1]に基づく計算機シミュレーションを行った。環境条件を Table 1 に示した。

残響時間を 0.2 s とし、音源との相対的位置は固定しながら 2 素子アレーの中心位置および向きをランダムに 5000 パターン変化させた。そして、比較的反射音に強いとされる PHAT 法[2,3]を用いて方向推定を行い、誤推定結果を分析した。Fig. 1 は、到来時間差 (TDOA) の誤推定結果を示すもので、+14 および -14 サンプルに多数の誤推定が発生した。Fig. 2 は誤推定が発生した場合の相互相関関数の例である。+18 サンプルの正解ピークのほか、-14、+14 サンプルの位置に高い相関値の虚ピーク①②が現れており、これらの虚ピークが誤推定の要因と考えられた。

次に、各 ch のインパルス応答 (Fig. 3) とこれらの虚ピークの関係を検討した。Fig. 3 より、直接音の直後に大きな反射音が到来しており、虚ピーク①の発生原因は ch 1 の直接音と ch 2 の初期反射音の相関、虚ピーク②の発生原因は初期反射音同士の相関であることがわかった。

このことから、反射音による推定性能劣化の主原因は初期反射音であり、音源方向推定の性能向上には初期反射音の抑圧が有効と考えられた。

3 提案法

3.1 初期反射音の抑圧

方向推定の前段において反射音除去を行う試みはこれまででもなされているが[4]、ここでは前節の分析結果より、特に初期反射音に着目した抑圧処理を加えることを提案する。提案法の構成を Fig. 4 に示す。初期反射音の抑圧は Fig. 4

Table 1 シミュレーション条件

マイクロホン数	2 素子
マイクロホン間隔	0.6 [m]
アレー配置の高さ	1.2 [m]
想定する部屋サイズ	9.0×5.0×2.4 [m]
目的音	男性音声
相対的音源位置	3 [m] 60 [°]
サンプリング周波数	12000 [Hz]

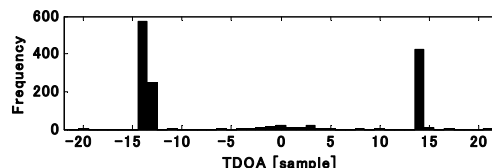


Fig. 1 誤推定方向の頻度分布

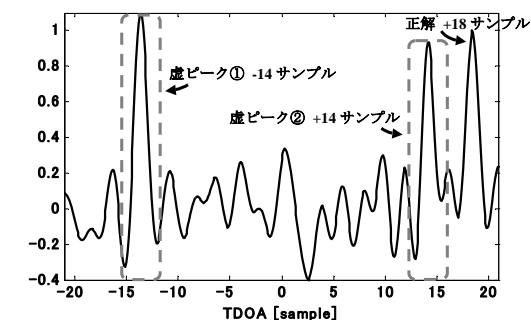


Fig. 2 PHAT 法による相関関数

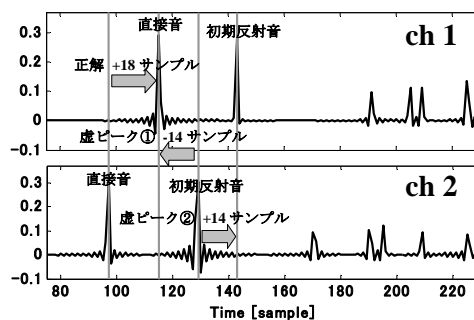


Fig. 3 各 ch のインパルス応答

に示すように各 ch 毎に線形予測フィルタによって行った。フィルタ係数 a_i ($i=1, \dots, N$) は出力 \hat{x}_M のパワーが最小となるように決定する。このときフィルタ長を初期反射音の長さ程度にすれば、初期反射音は抑圧できると考えられる。

3.2 相関関数の改善効果

初期反射音の抑圧処理を加えた提案法の有効性を検証した。前節 Fig. 2 の条件において、線形予測のフィルタ長 N を 500 とし PHAT 法を用

* DOA Estimation with Early Echo Reduction Using Linear Prediction, by UESUGI, Nobutoshi and KANEDA, Yutaka (Tokyo Denki Univ.).

い、虚ピークの抑圧効果を確認した。提案法の改善効果を Fig. 5 に示す。Fig. 5 が示すように、初期反射音を抑圧することで Fig. 2 で示した虚ピーク (± 14 サンプル付近) の高さを大幅に抑えられることがわかった。

3.3 フィルタ長

初期反射音の抑圧に適したフィルタ長 N の検討を行った。フィルタ長を 1~2000 の範囲で変化させ、最大虚ピークの相関値で抑圧効果を評価し Fig. 6 に示す。Fig. 6 が示すように、フィルタ長は 500 (約 40 ms) 程度で抑圧性能が飽和することがわかった。このことは、40 ms 程度の初期反射音を抑圧することは有効であるが、それ以降の反射音の抑圧は大きな効果を持たないことを示している。その理由としては、1) このフィルタでは中、後期反射音は抑圧量が小さい、2) 特定の到来方向を持たず多方向から到来する中、後期反射音は方向推定に与える影響は小さい、ことなどが考えられる。

3.4 提案法の性能評価

提案法の有効性を確認するため、シミュレーションにより音源方向の推定正解率を評価した。残響時間を 0~0.4 s の範囲で変化させた。初期反射音抑圧の線形予測フィルタ長は 500 とした。その他の条件は前節と同様である。

許容誤差を $\pm 5^\circ$ として求めた 2 つの従来法 (CC(相互相関関数)法, PHAT 法) と、それらに初期反射音抑圧 (EER : Early Echo Reduction) 処理を加えた提案法の推定正解率を Fig. 7 に示す。Fig. 7 が示すように、CC 法は残響時間が 0.05 s, PHAT 法は 0.15 s 程度まで良好に推定できたのに対し、EER 処理を加えた提案法では EER+CC, EER+PHAT 法ともに 0.25 s 程度まで良好に推定でき、本手法の有効性が確認された。

また EER 処理後では CC 法も PHAT 法とほぼ同様な性能となった。これは、EER 処理のための線形予測フィルタにより、初期反射音が抑圧されるとともに音声信号が白色化されたためと考えられる。

4 まとめ

本報告では、音源方向推定に及ぼす反射音影響を分析した結果、反射音による推定性能劣化の主原因は初期反射音とわかった。そこで、方向推定の前段に初期反射音抑圧を加える手法を提案し、大幅に性能向上することがわかった。本提案手法は、方向推定法に容易に適用可能の

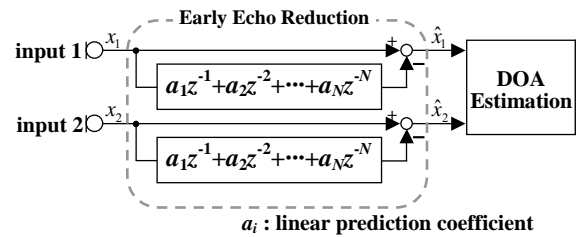


Fig. 4 提案法の構成

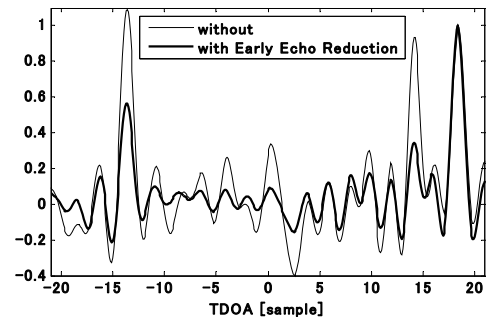


Fig. 5 提案法の改善効果

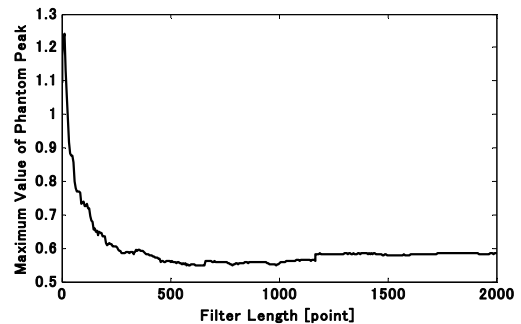


Fig. 6 フィルタ長と抑圧効果の関係

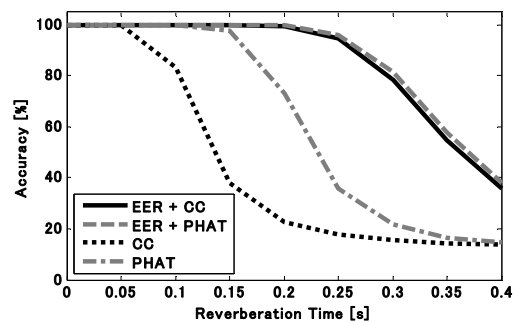


Fig. 7 初期反射音抑圧処理の性能評価

ため、多 ch アレー処理など他手法への適用も期待できる。

参考文献

- [1] J. B. Allen and D. A. Berkley, J. Acoust. Soc. Am., vol. 65, no. 4, pp. 943-950, Apr. 1979.
- [2] C. H. Knapp and G. C. Carter, IEEE Trans. ASSP, vol. 24, no. 4, pp. 320-327, Aug. 1976.
- [3] 上杉, 金田, 信学技報, EA2006-105, Jan. 2007.
- [4] 例えば, 大井, 陶山, 信学技報, EA2006-10, May 2006.