

空間エイリアシング回避による複数音源追尾の性能改善

A-8 Performance Improvement of Multiple Sound Source Tracking Using Spatial Aliasing Avoidance

中澤 克弥
Katsuya NAKAZAWA陶山 健仁
Kenji SUYAMA東京電機大学 工学部 電気電子工学科
School of Engineering, Tokyo Denki University

1 はじめに

マイクロホンアレーを用いた複数音源追尾手法として PAST-IPLS (Projection Approximation Subspace Tracking Interior Point Least Square) が提案されている [1]. 一般に, マイクロホン間隔を拡大すれば, 追尾精度向上が期待できる. 本研究では, マイクロホン間隔を拡大する. しかし, 高周波数帯域において, 空間エイリアシングによる誤推定の可能性がある. そこで, 全帯域の評価関数の乗算により空間エイリアシング回避を試みる. 実環境実験で検討手法の有効性を示す.

2 問題設定

時刻 n において, 音源信号 $s_i(n)$, $i = 1, 2$, を直線状に等間隔で配置した 4 個のマイクロホンで受音する. 受音信号は周波数領域で次式のようにベクトル表記される.

$$\mathbf{X}_t(k) = \sum_{i=1}^2 S_{i,t}(k) \mathbf{a}(\tau_{i,\theta(t)}) + \mathbf{\Gamma}_t(k) \quad (1)$$

ここで t はフレーム番号, k は周波数インデックス, $\mathbf{X}_t(k)$ は受音信号ベクトル, $S_{i,t}(k)$ は $s_i(n)$ の複素振幅, $\mathbf{a}(\tau_{i,\theta(t)})$ はステアリングベクトル, $\mathbf{\Gamma}_t(k)$ は雑音ベクトルである. 各マイクロホンの受音信号間には到達時間差 $\tau_{i,\theta(t)} = d \sin \theta_i(t)/c$ が生じる. ここで d はマイクロホン間隔, c は音速である. 複数音源追尾問題は受音信号 $\mathbf{X}_t(k)$ から到達時間差 $\tau_{i,\theta(t)}$ を推定する問題である.

3 PAST-IPLS による複数音源追尾

PAST-IPLS は信号部分空間と雑音部分空間の直交性を利用した手法である. PAST と IPLS の 2 つの逐次解法により構成され, 信号部分空間の算出と音源方向の探索を高速に行う. 図 1 に手法の流れ, 以下に手続きを示す.

1. $\mathbf{x}(n)$ を DFT (Discrete Fourier Transform) により周波数領域に変換する.
2. PAST で信号部分空間 $\mathbf{Q}_S(k, t)$ を更新する.
3. 各周波数帯域毎に次式の評価関数 $J_{I,k}(\tau)$ を算出する.

$$J_{I,k}(\tau) = \mathbf{a}_k^H(\tau) \mathbf{Q}_N(k, t) \mathbf{Q}_N^H(k, t) \mathbf{a}_k(\tau) \quad (2)$$

ここで, $\mathbf{Q}_N(k, t) \mathbf{Q}_N^H(k, t) = \mathbf{I} - \mathbf{Q}_S(k, t) \mathbf{Q}_S^H(k, t)$ である.

4. 全帯域の評価関数を乗算した $J_I(\tau)$ を算出する.
5. IPLS で評価関数 $J_I(\tau)$ を最小化する到達時間差 $\tau_{i,\theta(t)}$ を求める. PAST-IPLS の更新則は文献 [1] に従う.

$J_{I,k}(\tau)$ において, 音源方向に対応する局所解は全ての周波数で同じ方向に存在し, 音源方向でない局所解は周波数毎に異なる方向に存在する. そこで, 全周波数帯域の評価関数を乗算し, 音源方向以外の局所解を抑圧する.

複数音源追尾では, 一方の話者の非発話区間で, もう一方の音源を捉え, 再び発話区間に戻っても, 追尾対象が戻らない同一音源推定問題がある. そのため, 発話復帰時に IPLS の実行可能領域を再設定し, この問題を回避する. 提案法では遅延和アレー法と音源選択処理に基づき, 発話復帰時に IPLS の実行可能領域を再設定する.

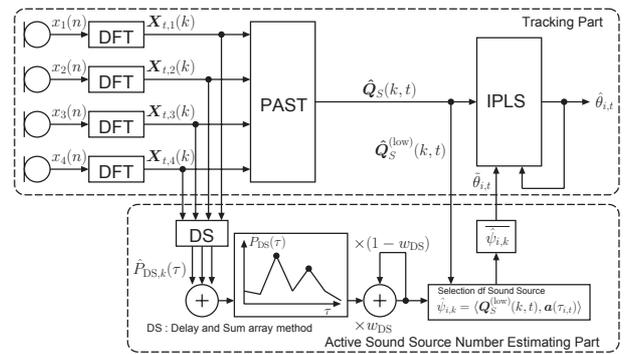


図 1 手法の流れ

4 実環境実験

提案法の有効性を示すため, 実環境実験を行った. マイクロホン間隔 $d = 0.16$ [m], サンプリング周波数 $f_s = 8$ [kHz], フレーム長 $N = 512$ とした. 4 種類の音声パターンに対して評価を行った. 評価指標として RMSE (Root Mean Square Error) と実時間処理の指標である RTF (Real Time Factor) を使用した. 表 1 に各パターンの RMSE を示す.

表 1 RMSE の結果

パターン	1	2	3	4
RMSE [°]	4.17	2.38	2.89	4.92

RTF はいずれも 0.058 であり, 実時間処理可能であった.

5 まとめ

本研究では, PAST-IPLS による複数音源追尾について検討した. 実環境実験より, 提案手法は高精度かつ実時間で複数音源追尾可能であることを示した.

参考文献

- [1] 大和田昇, 陶山健仁, “多峰性評価関数に対する逐次解法に基づく複数音源追尾,” 信学論 (A), Vol. J93-A, No.4, pp.248-259, April 2010.