

音源数未知の状況における全方位複数音源追尾

A-8

Omnidirectional Multiple Sound Source Tracking Under Unknown Number of Sound Sources

松永 拳
Ken MATSUNAGA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

近年、音声認識などの音声インターフェースの普及により、円形マイクロホンアレーを用いた全方位音源追尾 [1]~[2] が注目されている。文献 [1] では、パーティクルフィルタを用いて、全方位音源追尾に成功している。しかし、同一音源方向推定問題により、複数音源追尾が困難な場合がある。文献 [2] では、広間隔2マイクロホンペア毎の推定結果を統合したヒストグラムを作成し、全方位複数音源追尾に成功している。通常、音源方向推定のために、音源数を設定する必要がある。しかし、一般に音源数は未知である。そこで、本研究では、音源数未知の状態を想定した全方位複数音源追尾について検討する。

2 全方位音源追尾問題

円周上に $\alpha = 360^\circ/M$ 毎に配置した M 個のマイクロホンで、 N 個の移動音源信号、 $s_i(n)$, $i = 1, 2, \dots, N$, を受音する。ここで、 n は離散時間である。間隔が最大の2マイクロホンペアを用いた場合、マイクロホン間の到達時間差 $\tau_{p,i}(t)$, $p = 1, 2, \dots, M/2$, は次式となる。

$$\tau_{p,i}(t) = \frac{d_{pair} \cos(\theta_i(t) - (p-1)\alpha)}{c} \quad (1)$$

ここで、 c は音速、 d_{pair} は2マイクロホンの間隔である。全方位音源追尾問題は受音信号 $x_m(n)$, $m = 1, 2, \dots, M$, から各音源方向 $\theta_i(t)$, $i = 1, 2, \dots, N$, を推定する問題である。

3 提案法

図1に提案法の流れを示す。文献 [2] では、混合 wrapped コーシー分布を方向推定結果のヒストグラムに当てはめ、最尤推定を行い、複数音源追尾に成功している。混合 wrapped コーシー分布 $F(\theta)$ は次式で定義される。

$$F(\theta) = \sum_{i=1}^N \rho_i \left[\frac{1}{2\pi} \left(\frac{1 - \gamma_i^2}{1 + \gamma_i^2 - 2\gamma_i \cos(\theta - \mu_i)} \right) \right] \quad (2)$$

ここで、 N は混合数 (音源数)、 ρ_i は混合比、 μ_i は最頻値、 γ_i は半値半幅を表す。混合 wrapped コーシー分布を当てはめる際、音源数を設定する必要がある。しかし、音源数は未知である。そこで本研究では、音源数を過剰に設定し、最尤推定を行う。音源数を過剰に設定した場合、音源方向以外の結果も推定する可能性がある。誤った推定結果を取り除くために、各推定結果に対応した ρ_i を閾値で評価し、音源数選択を行う。ここで、全ての音

源が無音状態の場合、各推定結果の混合比は同一と考えられる。そこで、閾値は、 $C_{th} = 1/N$ と設定する。 C_{th} により、正しい推定結果のみを選択可能である。

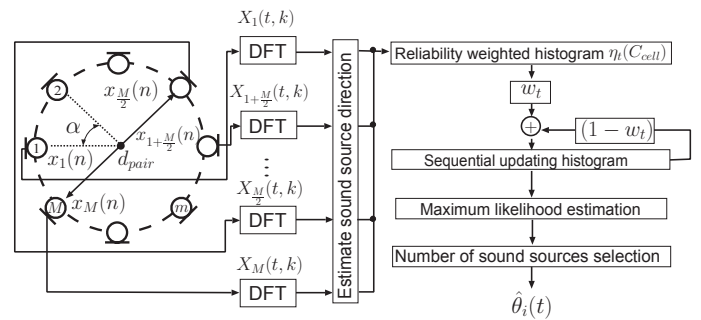


図1 提案法の流れ

4 実環境実験

提案法の有効性を示すために、実環境実験を行った。実験条件は、 $M=16$ 、サンプリング周波数を 8000[Hz]、フレーム長を 512、 $d_{pair}=0.3$ [m]、使用周波数帯域を 100[Hz]~ 2833[Hz]、過剰音源数を $N = 5$ 、信号長を 10[s] とした。音源は 12 パターンの異なる 2 つの音声の組み合わせを使用した。比較手法として、音源数を与えた ($N = 2$) 文献 [2] の手法を使用した。追尾精度の評価は、RMSE(Root Mean Square Error)、音源数選択の評価は、音源数正解率を用いた。12 パターンの平均 RMSE は、提案法が 4.30[°]、比較手法が 6.21[°] であった。また、12 パターンの平均音源数正解率は、88.8% であった。

5 まとめ

本研究では、音源数を過剰に設定し、最尤推定の結果から閾値を用いて音源数選択を行った。実環境実験より、提案法が音源数未知の状態でも高精度な追尾が可能であることを示した。

参考文献

- [1] I. Marković and I. Petrović, "Speaker localization and tracking with a microphone array on a mobile robot using Von Mises distribution and particle filtering", Robotics and Autonomous Systems, Vol.58, No.11, pp.1185-1196, Nov. 2010.
- [2] K. Matsunaga and K. Suyama, "Improvement of Omnidirectional Sound Source Tracking Performance By Expanding Microphone Width", Proceedings of ITC-CSCC 2018, pp.385-388, July 2018.