

## 探索状態を制御する 2 マイクロホン音源追尾手法

A-8 Sound Source Tracking via Two Microphones by Adjusting Search Ability

大宮 健太  
Kenta OMIYA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

マイクロホンによる複数音源追尾手法として、PSO(Particle Swarm Optimization)を用いた手法が提案されている [1]。PSO は、探索パラメータによって集中化能力と多様化能力を制御可能である。すなわち、追尾の状況に応じてパラメータ値を設定できれば、追尾性能の向上が期待できる。本研究では、分布当てはめに用いる混合コーシー分布のパラメータを利用し、探索パラメータの設定を試みる。実環境実験より提案法の有効性を示す。

## 2 問題設定

時刻  $n$  において 2 つの移動音源  $s_i(n)$ ,  $i = 1, 2$ , の信号を、間隔  $d$  で配置した 2 個のマイクロホンで受音するモデルを考える。マイクロホン間には到達時間差  $\tau_i(n) = d \sin \theta_i(n)/c$  が生じる。ここで、 $\theta_i(n)$  は音源方向、 $c$  は音速である。複数音源追尾問題は受音信号  $x_m(n)$ ,  $m = 1, 2$ , から  $\theta_i(n)$  を推定する問題である。

## 3 PSO

PSO は複数の個体と群から構成され、個体  $u$  は位置  $\mathbf{x}_u$  と速度  $\mathbf{v}_u$  をもつ。 $\mathbf{x}_u$  と  $\mathbf{v}_u$  は次式で更新する。

$$\mathbf{x}_u^{l+1} = \mathbf{x}_u^l + \mathbf{v}_u^{l+1} \quad (1)$$

$$\mathbf{v}_u^{l+1} = w\mathbf{v}_u^l + c_1 r_1 (\mathbf{p}_u^l - \mathbf{x}_u^l) + c_2 r_2 (\mathbf{p}_g^l - \mathbf{x}_u^l) \quad (2)$$

ここで、 $l$  は探索回数、 $w$  は慣性重み、 $r_1, r_2$  は  $[0, 1]$  の一様乱数、 $u$  番目の個体の  $\mathbf{p}_u^l$  は個体最良解、 $\mathbf{p}_g^l$  は群最良解、 $c_1$  は個体最良解方向への重み、 $c_2$  は群最良解方向への重みである。本研究では、MUSIC(Multiple Signal Classification) スペクトルを評価関数とし、MUSIC スペクトルのピークを PSO を用いて探索する。

## 4 提案法

図 1 に手法の流れを示す。

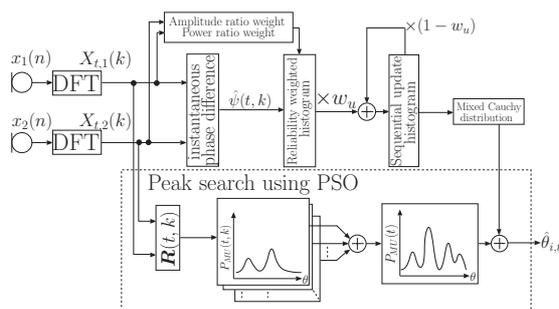


図 1 手法の流れ

$x_m(n)$  を DFT(Discrete Fourier Transform) で周波数領域に変換する。まず、低周波数帯域で瞬時位相差  $\hat{\psi}(t, k)$  を求め、逐次更新ヒストグラムを算出する。ここで、 $t$  はフレーム番号、 $k$  は周波数番号である。ヒストグラムに次式の混合コーシー分布  $F_{cd}(\theta, t)$  を当てはめ、PSO のペナルティ関数を算出する。

$$F_{cd}(\theta, t) = \sum_{i=1}^2 \rho_{i,t} \left[ \frac{1}{\pi} \left( \frac{\lambda_{i,t}}{(\theta - \mu_{i,t})^2 + \lambda_{i,t}^2} \right) \right] \quad (3)$$

ここで、 $\rho_{i,t}$  は混合比、 $\mu_{i,t}$  は最頻値、 $\lambda_{i,t}$  は半値半幅である。最後に、ペナルティ関数を高周波数帯域の加算 MUSIC スペクトルに対して付加し、PSO を用いて音源追尾する。

追尾状況に応じて、 $c_1, c_2$  を調整するため、 $\rho_{1,t}, \rho_{2,t}$  の大小関係を利用する。 $\rho_{1,t}$  が大きい場合、 $\theta_1(n)$  付近における MUSIC スペクトルのピークは鋭くなる。そのため、各個体が自由に探索できるように多様化能力を高める必要がある。そこで、 $c_1$  を大きくする。これに対して、 $\rho_{1,t}$  が小さい場合、 $\theta_2(n)$  付近における MUSIC スペクトルは鈍くなる。このとき、誤ったピークに停留する可能性があるため、集中化能力を高める必要がある。そこで、 $c_2$  を大きくする。 $\rho_{2,t}$  に対しても同様に行う。

## 5 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。実験ではサンプリング周波数を 8000[Hz]、フレーム長を 512、使用低周波数帯域を 500[Hz]~850[Hz]、使用高周波数帯域を 850[Hz]~4000[Hz]、信号長を 20[s]、 $d=0.2$ [m] とした。音源は異なる音声の組み合わせ 3 パターンを用いた。追尾精度の評価指標として RMSE(Root Mean Square Error) を用いた。比較手法には  $c_1 = 1.3, c_2 = 1.0$  に固定した場合を用いた。各パターンに対し 30 回の試行を行った。このときの各最良 RMSE の平均値は、提案法は  $5.21^\circ$ 、比較手法は  $6.81^\circ$  であった。これより、提案法は比較手法より追尾精度が高いことを確認した。

## 6 まとめ

本研究では、PSO の探索状態を制御するために、探索パラメータである  $c_1, c_2$  を、 $F_{cd}(\theta, t)$  より算出した  $\rho_{1,t}, \rho_{2,t}$  の大小関係を利用して設定した。実験結果より提案法の有効性を示した。

## 参考文献

[1] 平川真里, 陶山健仁, "PSO を用いた音声のスパース性に基づく複数音源追尾," 信学技報, CAS2013-76, pp.19-24, February 2014