非同期マイクロホンアレイを搭載した 複数ロボットによる音環境マップの協調オンライン推定

Online Localization of Multiple Sound Sources and Multiple Robots with Asynchronous Microphone Arrays

○関口 航平 (京大)　坂東 宜昭 (京大)　中村 圭佑 (HRI)
中臺一博 (HRI)　糸山 克寿 (京大)　吉井和佳 (京大)

Kouhei Sekiguchi[1], Yoshiaki bando[1], Keisuke Nakamura[2]
Kazuhiro Nakadai[2], Katsutoshi Itoyama[1], Kazuyoshi Yoshii[1]
[1]Kyoto University, [2]Honda Research Institute Japan

This paper presents an online method for localizing the positions of multiple sound sources and stationary robots and synchronizing microphone arrays attached to those robots. Since each robot can estimate only the directions of sound sources, the two-dimensional source positions can be estimated from the source direction estimated by each robot using a triangulation. In addition, mixture signals can be separated accurately by regarding multiple microphone arrays as one big array. To perform these tasks, some methods have been proposed for localizing and synchronizing microphone arrays. These methods, however, assume only a single sound source exists. To overcome this limitation, we estimate the directions of arrival (DOAs) and separate observed signals to estimate the time differences of arrival (TDOAs) by using microphone array techniques, and integrate the DOAs and TDOAs by using a state-space model. The latent variables are estimated in an online manner with a FastSLAM2.0 algorithm.

Key Words: Robot audition

1　はじめに

音源定位や分離といったマイクロホンアレイ処理を用いて，ロボット周辺の音環境理解を行う研究がなされている [1]．マイクロホンアレイを搭載した１台のロボットを用いた場合には，音源の方向しか推定することができないが，複数台のロボットを用いることにより２次元平面上での音源位置を推定することが可能となる [2,3]．また，複数台のロボット全体を一つの大きなマイクロホンアレイとみなして音源分離を行うことにより [4]，雑音が多く存在するような厳しい環境においても頑健な音源分離が実現できると期待される．

複数台のロボットを用いてマイクロホンアレイ処理を行うためには，各ロボットの位置とロボット間でのマイクロホンアレイの同期が必要となる．この問題を解決するために，非同期の複数マイクロホンを用いて，各マイクロホンの位置と周囲の音源位置，マイクロホン間の同期時刻ずれを同時に推定する研究がなされてきた [5–7]．これらの手法は音源が１つであるという制約があるため，複数音源存在下で用いることはできない．

本研究では，マイクロホンアレイを搭載した複数台のロボットを用いて，複数音源存在下で各ロボットの位置と向き，各音源の位置，ロボット間の同期時刻ずれを同時推定する手法の開発を行う．これらを推定するために，各音源の到来方向とロボット間での到達時間差を推定する事が必要となる．各音源の到来方向を推定することは従来法を用いることにより可能だが，観測される混合音から各音源の到達時間差を直接推定することは困難である．そこで，各ロボットごとに音源到来方向の情報をもとにマイクロホンアレイ処理を用いて音源分離を行い，分離音から到達時間差の推定を行う．推定された到来方向と到達時間差を状態空間モデルを用いて統合し，FastSLAM2.0 によりロボットの位置と向き，音源の位置，同期時刻ずれの推定を行う．

2　非同期複数マイクロホンアレイ位置・同期時刻ずれのオンライン推定

M チャンネルマイクロホンアレイを搭載した I 台のロボットが静止し，音源が複数存在する状況において，二次元平面上での各ロボットの位置・向き $(r_i^x, r_i^y, r_i^\theta\ i = 1, \cdots, I)$，各音源の位置 $(s_n^x, s_n^y\ n = 1, \cdots, N)$，ロボット 1 を基準としたときのロボット j の同期時刻ずれ $(\xi_j\ j = 2, \cdots, R)$ を推定する．状態の推定には，各ロボットから見た音源方向，観測音のロボット間の到達時間差を用いる．ただし，どの観測も部屋の残響の影響などによりノイズを含んでいる．状態空間モデルを用いてモデル化し，FastSLAM2.0 [8] を用いることで，ノイズに頑健な状態の推定を行う．

入力	M チャンネルマイクロホンアレイを搭載した I 台のロボットでの観測音
出力	(1) ロボット i の状態 (位置，向き) $(r_i^x, r_i^y, r_i^\theta)$ (2) 音源 n の位置 (s_n^x, s_n^y) (3) ロボット 1 と j の同期時刻ずれ ξ_j
仮定	(1) 各ロボットは静止 (2) 複数音源のうち少なくもと 1 つは移動音源

2.1　観測

ロボット・音源位置，同期時刻ずれの推定は音源到来方向と到達時間差を用いる．入力された観測音からこれらを推定する方法について述べる．

2.1.1　音源到来方向

各ロボットでそれぞれ同期録音した観測音に対して Multiple Signal Classification (MUSIC) 法 [9] を用いることで，音源到来方向の推定を行う．ただし，得られる角度はロボットの正面を基準とした角度である．MUSIC 法は複数音源が存在する場合でも各音源の到来方向を推定できるが，事前に音源数を指定する必

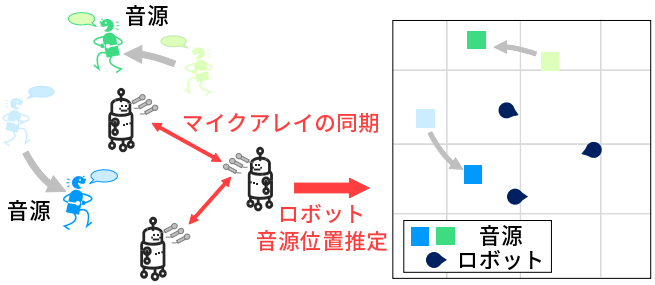


Fig.1 Localizing the positions of the sound sources and robots and synchronizing multiple microphone arrays.

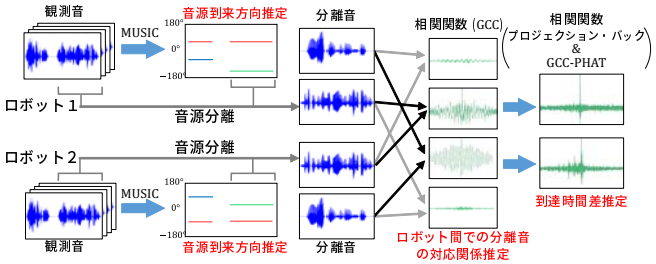


Fig.2 How to estimate DOAs and TDOAs when there are multiple sound sources.

要がある．実際の環境では音源数は時変であるため，適切な値を設定することは困難であるが，実際の音源数と指定した音源数が異なる場合でも，推定精度が低下する可能性はあるが，推定が必ずしも失敗するわけではない．

2.1.2 到達時間差

到達時間差とは，ある音源が各マイクロホンにどれだけ遅れて到達したかということを表している．各ロボットのマイクロホンアレイが同期できている場合，到達時間差からロボット間での音源との距離差を得ることができる．実際にはマイクロホンアレイは同期しておらず，同期時刻ずれがあるため，到達時間差は距離差と同期時刻ずれの情報の混合となっている．本稿では，ロボット 1 を基準として到達時間差を計算する．

到達時間差の推定は，音源が一つしか存在しない場合，次のように行う．まず，観測音の相関関数を Generalized Cross Correlation with Phase Transform (GCC-PHAT) [10] を用いて以下のように計算する．

$$G_{PHAT}(f) = \frac{X_{m_1}(f)X^*_{m_2}(f)}{|X_{m_1}(f)X^*_{m_2}(f)|} \tag{1}$$

ここで，X_{m_i} はマイク m_i での観測音のフーリエ変換を表す．次に，相関関数を逆フーリエ変換を用いて時間領域に変換し，その最大値が閾値以上であれば，最大値を取る時間ずれを到達時間差とする．従って，到達時間差 ξ は以下の式で計算される．

$$\xi = \mathrm{argmax}_\xi \int G_{PHAT}(f)e^{j2\pi f\xi}df \tag{2}$$

音源が複数存在する場合には，各音源の到達時間差を推定し，到達時間差と音源到来方向の対応関係を推定する必要があるが，上記の方法ではこれを行うことができない．なぜなら，混合音の相関関数を計算した場合，最大値が閾値を超える可能性はあるが，その到達時間差がどの音源到来方向に対応しているのかを推定することが困難なためである．

この問題を解決するために，マイクロホンアレイ処理を用いて観測された混合音を各音源の信号に分離を行う．音源分離の際には音源到来方向を用いるため，各分離音と音源到来方向の対応関係は分かる．到達時間差を計算する際に基準とするロボットの分離音が，他のロボットのどの分離音に対応するかを推定するために，Generalized Cross Correlation (GCC) [10] を用いて相関関数を計算する．ロボット i が l 番目に見つけた音源の分離音のフーリエ変換を $Y_{i,l}$ と定めると，Y_{i_1,l_1} と Y_{i_2,l_2} の GCC による相関関数 $G_{i_1l_1,i_2l_2}$ は次のようになる．

$$G_{i_1l_1,i_2l_2}(f) = Y_{i_1l_1}(f)Y^*_{i_2l_2}(f). \tag{3}$$

式 (1) と式 (3) から分かるように，GCC-PHAT は信号の位相にのみ注目しているのに対し，GCC は信号のパワーと位相の両方に注目しているという点で，これらは異なる．$G_{i_1l_1,i_2l_2}$ の最大値が閾値以上である場合，Y_{i_1,l_1} と Y_{i_2,l_2} は同一音源の分離音であるとみなす．

分離音から到達時間差を推定する際に問題となるのは，音源分離により観測音の位相がずれてしまうため，分離音から推定された到達時間差は実際の到達時間差と異なるということである．この問題を解決するために，分離行列の逆行列を掛けて位相を元に戻せば良い．ある時刻 k での音源数を N_k とし，ロボット i の m 番目のマイクでの観測音のフーリエ変換を $X_{i,m}$ と定める．音源 n のマイク m での位相を復元するためには，分離音 $Y_{i,n}$ に対して分離行列の逆行列の (m,n) 成分を掛ければ良い．この処理はプロジェクションバック [11] と呼ばれ，本来はブラインド音源分離手法のスケーリング問題を解決するために用いられていた．

2.2 状態空間モデル

音源到来方向と到達時間差からロボット状態 (位置・向き)，音源位置，同期時刻ずれを推定するために，状態空間モデルを用いてモデル化を行う．状態空間モデルの潜在状態 $\boldsymbol{z}_t$ は，ロボットの位置・向き，音源位置，同期時刻ずれに加えて，音源の移動先を予測するために音源の速度と進行方向 (s^v, s^θ) を追加した $4I + 3N - 1$ 次元のベクトルとして以下のように定義する．

$$\boldsymbol{z}_k = [\boldsymbol{r}_1, \cdots, \boldsymbol{r}_I, \boldsymbol{s}_{1,k}, \cdots, \boldsymbol{s}_{N,k}, \boldsymbol{\tau}] \tag{4}$$

ここで，$\boldsymbol{r}_i$，$\boldsymbol{s}_{k,n}$，$\boldsymbol{\tau}$ は以下のように定義される．

$$\boldsymbol{r}_i = [r^x_i, r^y_i, r^\theta_i] \tag{5}$$

$$\boldsymbol{s}_{k,n} = [s^x_{k,n}, s^y_{k,n}, s^v_{k,n}, s^\theta_{k,n}] \tag{6}$$

$$\boldsymbol{\tau} = [\tau_{12}, \cdots, \tau_{1I}]. \tag{7}$$

2.2.1 状態遷移モデル

ロボットは静止しているため，状態遷移は音源についてのみ行う．音源は直進運動モデルに従って移動すると仮定するが，実際にはどのように動くかわからず，速度も変化するため，各状態がガウス分布に従うと仮定する．従って，状態遷移は次の式で表される．

$$p(\boldsymbol{s}_{k+1,n}|\boldsymbol{s}_{k,n}) = \mathcal{N}\left(\begin{bmatrix} s^x_{k,n} + s^v_{k,n}\cos(s^\theta_{k,n})\Delta t \\ s^y_{k,n} + s^v_{k,n}\sin(s^\theta_{k,n})\Delta t \\ s^v_{k,n} \\ s^\theta_{k,n} \end{bmatrix}, Q\right), \tag{8}$$

ここで，$Q \in \mathbb{R}^{3N\times 3N}$ はモデル誤差を表す共分散行列，Δt は 1 つ前の観測からの経過時間である．

2.2.2 観測モデル

観測は各ロボットから見た音源到来方向と到達時間差である．全ての観測は独立であるため，観測モデル $p(\boldsymbol{\phi}_k, \boldsymbol{\xi}_k|\boldsymbol{s}_k, \boldsymbol{x}, \boldsymbol{\tau})$ は次のように計算される．

$$p(\boldsymbol{\phi}_k, \boldsymbol{\xi}_k|\boldsymbol{s}_k, \boldsymbol{x}, \boldsymbol{\tau}) = \prod_{n=1}^{N}\left(\prod_{i=1}^{I} p(\phi_{k,i,n}|\boldsymbol{s}_k, \boldsymbol{x}) \prod_{j=2}^{I} p(\xi_{k,j,n}|\boldsymbol{s}_k, \boldsymbol{x}, \boldsymbol{\tau})\right). \tag{9}$$

$\phi_{k,i,n}$ を時刻 k でロボット i から見た音源 n の方向，$\xi_{k,j,n}$ を時刻 k での音源 n のロボット 1 と j の間の到達時間差と定める．各観測がガウス分布に従って分布すると仮定すると，$p(\phi_{k,i,n}|\boldsymbol{s}_k, \boldsymbol{x})$ と $p(\xi_{k,j,n}|\boldsymbol{s}_k, \boldsymbol{x})$ は以下のように表される．

$$p(\phi_{k,i,n}|\boldsymbol{s}_k, \boldsymbol{x}) = \mathcal{N}\left(\arctan\left(\frac{s^y_{k,n} - r^y_{k,i}}{s^x_{k,n} - r^x_{k,i}}\right) - r^\theta_{k,i}, \sigma^2_\phi\right) \tag{10}$$

$$p(\xi_{k,j,n}|\boldsymbol{s}_k, \boldsymbol{x}, \boldsymbol{\tau}) = \mathcal{N}\left((l_{k,j,n} - l_{k,1,n})/C - \tau_{1j}, \sigma^2_\xi\right) \tag{11}$$

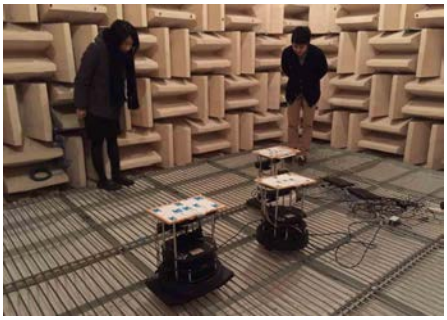
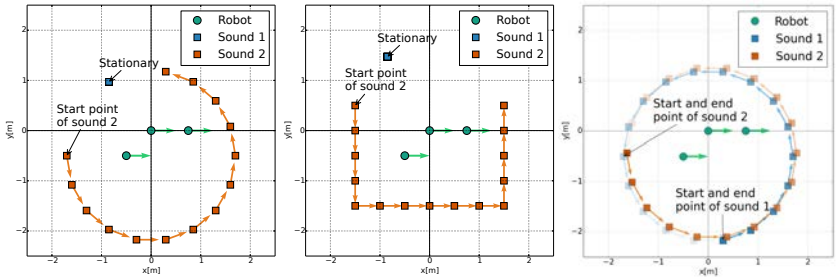


Fig.3 Experimental condition in an anechoic chamber. There are two sound sources (people) and three robots with 8-ch microphone arrays.



(a) パターン 1　(b) パターン 2　(c) パターン 3

Fig.4 Configuration of robot positions and sound source movements. Squares and circles indicate the sound source and robot positions, respectively. Arrows on squares indicate the movement directions of sound sources, and arrows on circles indicate the directions of robots.

ここで，σ_ϕ^2 と σ_τ^2 は各観測の分散パラメータ，C は音速，$l_{k,i,n}$ はロボット i と音源 n の距離を表し，$l_{k,i,n} = \sqrt{(s^x_{k,n} - r^x_{k,i})^2 + (s^y_{k,n} - r^y_{k,i})^2}$ である．

2.3 状態推定アルゴリズム

ロボット状態，音源の位置，同期時刻ずれの推定は FastSLAM2.0 [8] を用いて行う．FastSLAM2.0 は，ロボットが移動して周辺地図を推定する通常の SLAM 問題を解くために開発された手法である．通常の SLAM でのロボット状態が本研究での音源位置に，周辺地図がロボット状態と同期時刻ずれに対応するとみなすことで，FastSLAM2.0 を用いて状態の推定を行うことができる．ここでは，FastSLAM2.0 の概要と，一般的な SLAM 問題に適応する場合との違いについて説明する．

FastSLAM2.0 はサンプルの集合を用いて事後分布 $p(\boldsymbol{s}_k, \boldsymbol{r}, \boldsymbol{\tau} | \boldsymbol{\phi}_{1:k}, \boldsymbol{\xi}_{1:k})$ を近似する．同時に複数の音源の観測が得られた場合には，2つ目以降の観測の状態遷移時には Δt を 0 として，逐次的に更新を行う．まず，時刻 k でのサンプル u に対して，観測が既知の音源のどれかから生成されたのか，新たな音源から生成されたのかを判定する．これは，ある既知の音源から生成されたと仮定して尤度を計算し，尤度の最大値が閾値以下であれば新たな音源から生成されたとみなし，閾値以上であれば尤度が最大となる音源から生成されたとみなす．次に，既知の音源から生成された場合には，ロボット状態と同期時刻ずれを extended Kalman filter (EKF) を用いて更新する．新たな音源から生成された場合には，ロボット状態と同期時刻ずれの更新は行わず，新たな音源の位置のみを三角測量を用いて決定する．ロボット状態の不確実性と音源到来方向の誤差により，三角測量の交点は最大で ${}_IC_2$ 通り存在する．そのため，新しい音源位置はこれらの交点の平均として定め，以下のように計算する．

$$\begin{bmatrix} s^{[u]x}_{k,\mathrm{new}} \\ s^{[u]y}_{k,\mathrm{new}} \end{bmatrix} = \frac{1}{{}_IC_2} \sum_{r_1}^{I} \sum_{r_2 \neq r_1}^{I} \begin{bmatrix} \frac{\alpha^{[u]}_{k,r_1} - \alpha^{[u]}_{k,r_2}}{\tan\psi^{[u]}_{k,r_2} - \tan\psi^{[u]}_{k,r_1}} \\ \frac{\alpha^{[u]}_{k,r_1}\tan\psi^{[u]}_{k,r_2} - \alpha^{[u]}_{k,r_2}\tan\psi^{[u]}_{k,r_1}}{\tan\psi^{[u]}_{k,r_2} - \tan\psi^{[u]}_{k,r_1}} \end{bmatrix}, \tag{12}$$

ここで，$\psi^{[u]}_{k,r_i}$ と $\alpha^{[u]}_{k,r_i}$ は以下のように定義した．

$$\psi^{[u]}_{k,r_i} = r^{[u]\theta}_{k-1,r_i} + \phi_{k,r_i,\mathrm{new}} \tag{13}$$

$$\alpha^{[u]}_{k,r_i} = r^{[u]y}_{k-1,r_i} - \tan(\psi^{[u]}_{k,r_i}) r^{[u]x}_{k-1,r_i} \tag{14}$$

ロボット状態と同期時刻ずれの最終的な推定結果は各パーティクルの重み付き平均として計算した．音源については，パーティクルごとに音源数が異なる場合があるため，重み付き平均を計算することができず，以下のように推定結果を求めた．まず各パーティクルの音源位置の推定結果を，ロボットの重心から見た方向を用いて，K-means 法を用いてクラスタリングした．パラメータ K は各パーティクルの音源数の重み付き平均を切り上げたものとした．次に，各クラスタについて重み付き平均を計算して，K 個の推定結果を出力する．

3 評価実験

3台のロボットと2つの音源を用いて，提案法を評価するための実験を行った．

3.1 実験設定

8チャンネルマイクロホンアレイを搭載した3台のロボットと2つの音源を用いて，無響室で実験を行った (図 3)．音源の移動については以下の3つのパターンで実験を行った (図 4)．

1. **パターン 1:** 一つの音源が静止し，もう一つの音源が円周上を移動．録音時間は 40 秒．
2. **パターン 2:** 一つの音源はパターン 1 より 50cm だけ離れた位置に静止し，もう一つの音源は四角形の上を移動．録音時間は 45 秒．
3. **パターン 3:** 2つの音源が同一の円周上を異なる位置を始点として移動．録音時間は５５秒．

非同期で録音した場合，正しい同期時刻ずれが分からないため，マイクロホンアレイ同士を有線で接続し，同期録音を行った後，ロボット 2, 3 の観測音をそれぞれ 10ms, 5ms ずつ意図的にずらした．

FastSLAM2.0 アルゴリズムのパラメータは以下のように定めた．各時刻でのパーティクル数は 50000，各パーティクルの初期化はランダムに行った．音源到来方向の標準偏差は 5°，到達時間差の標準偏差は 0.1ms とした．その他のパラメータについては実験的に決定した．音源到来方向の推定と音源分離はオープンソース・ソフトウェア *HARK* に実装されている MUSIC と geometric high-order dicorrelation-based source separation (GHDSS) を用いた．同じ観測で複数回状態を更新してしまうのを避けるために，状態の更新は直近 2 s 以内に現れていない観測が得られた場合のみ行った．

ロボット状態，同期時刻ずれの評価は正解との誤差の平均として計算した．音源については，推定結果と正解音源の対応関係がわからず，推定された音源数は実際の音源数と異なる場合があるため，以下のようにして推定誤差を計算する．まず，全ての推定結果に対して，最も近い正解音源のラベルを割り当てる．ある正解音源の推定誤差は，その音源のラベルを持つ推定結果が存在する場合には，そのうちの最も近い推定結果との誤差として計算し，ラベルを持つ推定結果が存在しない場合には計算しないこととした．

3.2 実験結果

図 6 にロボットの位置と向き，同期時刻ずれの推定誤差を示す．全てのパターンにおいて，高精度に推定できていることが確認でき，最終的な推定誤差はロボット位置が 0.05 m 以下，ロボットの向きが 10° 以下，同期時刻ずれが 0.2 ms 以下となっていた．録音のサンプリングレートは 16 kHz であったため，0.2 ms の同期時刻ずれは 3.2 サンプルに相当し，これは音源分離を行う際に問題にならない程度の誤差であるといえる．

図**??**に音源位置の推定誤差と音源数の推定結果を示す．音源数の推定結果は各パーティクルの重み付き平均として計算した．パ

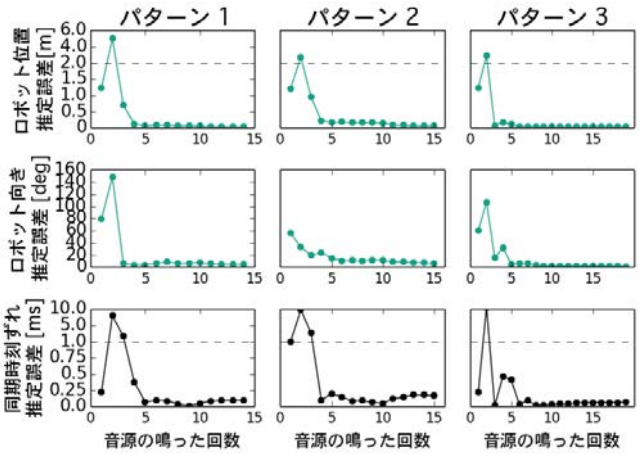


Fig.5 Estimation errors of the robot positions, the robot angles, and the time offsets.

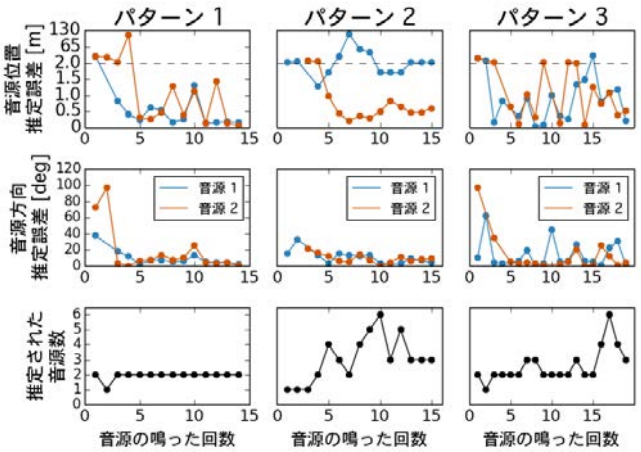


Fig.6 Estimation errors of the sound source positions and the sound source directions viewed from the centroid of the robot positions, and the estimated number of sound sources.

ターン 1 では，最終的な推定誤差はどちらの音源も 15 cm 以下であり，音源数も殆どの場合において正しい値を取った．パターン 2 では，静止した音源の推定誤差がとても大きくなっているが，推定された音源の方向は正解音源の方向と近くなっている．

幾つかの場合において音源位置の推定に失敗した理由として考えられるのは，ロボット間の距離と比較してロボットと音源の距離が長く，音源到来方向推定の少しの誤差が音源位置の大きな推定誤差につながったためだと考えられる．図 7 は１４回目の観測が得られた後の各パーティクルの音源の推定結果を表している．音源の推定結果，つまり各パーティクルの重み付き平均は正解位置から離れているが，ロボットから見た音源の方向は概ね正しく推定できていることが確認できる．

この問題を解決するための方法の一つは，音源の推定結果の分散が小さくなるようにロボットを移動させることである．このような手段は，アクティブオーディションの研究で取り組まれている [12]．この手法を拡張することで本研究に適用できると考えられる．

4　おわりに

本稿では，複数音源存在下でマイクロホンアレイを搭載した各ロボットの位置と向き，音源位置，ロボット間でのマイクロホンアレイの同期時刻ずれをオンラインで同時推定する手法を開発した．非同期な複数マイクロホンを用いた従来の研究では，音源数が一つであることを仮定していたが，本研究では音源到来方向推定や音源分離といったマイクロホンアレイ処理を用いること

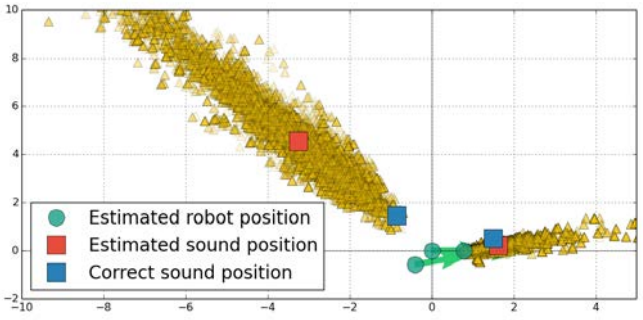


Fig.7 The experimental result in the Pattern 2 after the 14-th measurement. Blue squares, yellow triangles, and red squares indicate the correct sound source positions, the sound source positions of each particle, and the weighted mean of the yellow triangles, respectively.

で，観測される混合音から各音源の到来方向と到達時間差の推定を行った．推定された音源到来方向と到達時間差は状態空間モデルを用いて統合され，FastSLAM という手法により各状態の推定を行った．無響室での実験では，全てのパターンにおいてロボットの位置と向き，同期ずれの推定誤差がそれぞれ 5cm，10°，0.2ms 以下となった．一方，音源の推定はいくつかの場合で失敗したが，ロボットの重心から見た推定結果の方向は概ね正しく推定できていることが確認できた．今後は，ロボットを音源の推定結果の分散が小さくなるように移動させ，音源推定結果の向上を行う予定である．

参考文献

[1] H. G. Okuno et al. Robot audition: Missing feature theory approach and active audition. In *Robotics Research*, volume 70, pages 227–244. Springer, 2011.

[2] T. Nakashima et al. *Natural Interaction with Robots, Knowbots and Smartphones*, chapter Integration of Multiple Sound Source Localization Results for Speaker Identification in Multiparty Dialogue System, pages 153–165. Springer, 2014.

[3] E. Martinson et al. Optimizing a reconfigurable robotic microphone array. In *IEEE/RSJ IROS*, pages 125–130, 2011.

[4] K. Sekiguchi et al. Optimizing the layout of multiple mobile robots for cooperative sound source separation. In *IEEE/RSJ IROS*, pages 5548–5554, 2015.

[5] D. Su et al. Simultaneous asynchronous microphone array calibration and sound source localisation. In *IEEE/RSJ IROS*, pages 5561–5567, 2015.

[6] N. Ono et al. Blind alignment of asynchronously recorded signals for distributed microphone array. In *WASPAA*, pages 161–164, 2009.

[7] H. Miura et al. SLAM-based online calibration of asynchronous microphone array for robot audition. In *IEEE/RSJ IROS*, pages 524–529, 2011.

[8] S. Thrun et al. FASTSLAM: An efficient solution to the simultaneous localization and mapping problem with unknown data association. *J. Machine Learning Research*, 2004.

[9] R. Schmidt et al. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas and Propagation*, 34(3):276–280, 1986.

[10] C. Knapp and G. Carter. The generalized correlation method for estimation of time delay. *IEEE Trans. Acoustics, Speech and Signal Processing*, 24(4):320–327, 1976.

[11] N. Murata et al. An approach to blind source separation based on temporal structure of speech signals. *Neurocomputing*, 41(1-4):1–24, 2001.

[12] E. Vincent et al. Audio source localization by optimal control of a mobile robot. In *IEEE ICASSP*, pages 5630–5634, 2015.