

複数マイクロホンアレイの同期および位置・姿勢推定の同時最適化の検討

杉山 地塩¹, 糸山 克寿¹, 西田 健次¹, 中臺一博^{1,2}

1 東京工業大学 工学院 システム制御系 2 (株) ホンダ・リサーチ・インスティテュート・ジャパン

1 はじめに

複数のマイクロホンから構成されるマイクロホンアレイを用いて、観測信号から音源の方向を推定する音源定位は、音響信号処理の基本的な技術として広く研究が行われている [1]. 1つのマイクロホンアレイでは、音源の方向情報しか得られないことから、近年では複数のマイクロホンアレイを用いて、音源の位置を推定する手法の研究も盛んに行われている [2, 3]. 一方で、複数のマイクロホンを使う場合は、マイクロホンアレイの位置、向き、時刻同期といったパラメータの校正が必要であり、これらを自動校正しようとする試みも行われているが、限定的な校正にとどまっていた [4, 5]. また、探査空間が大きくなると局所最適解に陥りやすくなるという問題もある. そこで、本稿では向きを含む定位について、マイクロホンアレイ間の同期を仮定しない場合にも局所最適解に陥りにくい推定手法の構築を目指す.

2 提案手法

本稿では、 M 個の音源から発した音響信号を N 個のマイクロホンアレイで観測し、音源位置およびマイクロホンアレイの位置・回転・観測時刻のオフセットを推定する問題について考える.

図 1 に、音源方向と観測時間差のモデルを示す. 音源 i からの音響信号を観測し、マイクロホンアレイ j から見た音源方向の単位ベクトル $\mathbf{d}_{i,j}$ とマイクロホンアレイ j, k 間の観測時間差 $\tau_{i,(j,k)}$ が得られたとすれば、これらにをもとに位置・向き・時間オフセットを最適化する目的関数は式 (1), (2) となる [6].

ただし、音源位置 $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_M)$ 、マイクロホンアレイ位置 $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_N)$ 、向き $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)$ 、観測時刻オフセット $\boldsymbol{\delta} = (\delta_1, \dots, \delta_N)$ であり、 \mathbf{R}_j は θ_j に対応する回転行列を表す.

ここで、式 (1) では定位のスケールが一意に定まらない. よって、音源方向に基づく推定値 $\tilde{\mathbf{S}}, \tilde{\mathbf{A}}$ を同じ割合 α で拡張し、システム全体の推定値 $\hat{\mathbf{S}} = \alpha \tilde{\mathbf{S}}, \hat{\mathbf{A}} = \alpha \tilde{\mathbf{A}}$ とする.

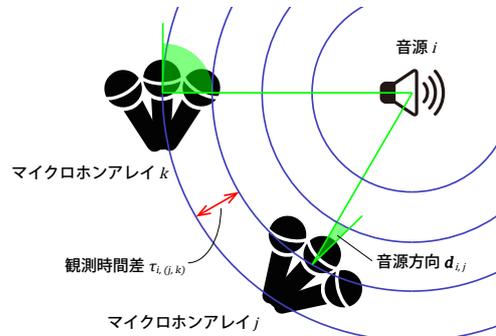


Fig. 1 音源方向・観測時間差のモデル

$$\tilde{\mathbf{S}}, \tilde{\mathbf{A}}, \hat{\boldsymbol{\theta}} = \underset{\mathbf{S}, \mathbf{A}, \boldsymbol{\theta}}{\operatorname{argmin}} \sum_{i=1}^M \sum_{j=1}^N D_{i,j}$$

$$D_{i,j} = \left\| \mathbf{R}_j^T (\mathbf{s}_i - \mathbf{a}_j) - \mathbf{d}_{i,j} \right\|_2 \left\| \mathbf{s}_i - \mathbf{a}_j \right\|_2 \quad (1)$$

$$\hat{\mathbf{S}}, \hat{\mathbf{A}}, \hat{\boldsymbol{\delta}} = \underset{\mathbf{S}, \mathbf{A}, \boldsymbol{\delta}}{\operatorname{argmin}} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^N T_{i,j,k}$$

$$T_{i,j,k} = \left\| \mathbf{s}_i - \mathbf{a}_j \right\|_2 - \left\| \mathbf{s}_i - \mathbf{a}_k \right\|_2 + c \delta_j - c \delta_k - c \tau_{i,(j,k)} \quad (j < k) \quad (2)$$

従来法 [6] では式 (1) を解いた後に観測時間差から座標のスケールを求める. しかし、音源方向を用いた最適化は位置のスケールが推定できないだけでなく、局所最適解に陥りやすい. そこで、第一の提案手法として式 (1), (2) を交互に計算し最適化を行う手法 (以下「繰り返し」「iterate」) を挙げる.

さらに、音源方向・観測時間差についての目的関数を同時に用いる手法として、最小化する目的関数 D, T の和をとることで変数すべてを同時に最適化する手法 (以下「連立」「simultaneous」), 目的関数 $D_{i,j}, T_{i,j,k}$ から共通項 $\left\| \mathbf{s}_i - \mathbf{a}_j \right\|_2$ を消去して評価する手法 (以下「統合」「integrated」) を提案し、その精度を比較する. 連立・統合法は式 (3), (4) によって定義される.

$$\hat{\mathbf{S}}, \hat{\mathbf{A}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}} = \underset{\mathbf{S}, \mathbf{A}, \boldsymbol{\theta}, \boldsymbol{\delta}}{\operatorname{argmin}} \sum_{i=1}^M \sum_{j=1}^N \left(D_{i,j} + \sum_{k=1}^N T_{i,j,k} \right)$$

$$(j < k) \quad (3)$$

Simultaneous Optimization of Synchronization and Position and Rotation Estimation of Multiple Microphone Arrays
Chishio Sugiyama¹, Katsutoshi Itoyama¹, Kenji Nishida¹, Kazuhiro Nakadai^{1,2}

¹ Dept. of Systems and Control Engineering, School of Engineering, Tokyo Institute of Technology

² Honda Research Institute Japan Co., Ltd.

$$\hat{\mathbf{S}}, \hat{\mathbf{A}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\delta}} = \underset{\mathbf{S}, \mathbf{A}, \boldsymbol{\theta}, \boldsymbol{\delta}}{\operatorname{argmin}} \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^N I_{i,j,k}$$

$$I_{i,j,k} = \left\| \mathbf{R}_j^T (\mathbf{s}_i - \mathbf{a}_j) - \mathbf{d}_{i,j} \left(T_{i,j,k} - \|\mathbf{s}_i - \mathbf{a}_j\|_2 \right) \right\|_2^2 \quad (j \leq k) \quad (4)$$

本稿では各最適化計算にあたり、内点法を用いて解の探索を行う。

3 実験・考察

提案法の有効性を検証するため、2次元空間における数値シミュレーションを行った。120 × 120[m]の空間に一様分布の乱数で音源7個、マイクロホンアレイ3個を配置し、マイクロホンアレイの向きは一様分布、観測時刻のオフセットは $\mathcal{N}(0, 10^{-4})$ [s]の分布の乱数で100通り生成した。音源数は変数と目的関数の個数の関係から、 $N = 3$ のとき式(1)が理論上解けるように設定した。また、マイクロホンアレイのうち一つの座標を(0, 0)、回転角と観測時刻のオフセットをそれぞれ0[deg], 0[s]と固定した。

各音源から独立した音響信号を1回ずつ観測したと仮定し、マイクロホンアレイに対する音源方向 \mathbf{d} 、マイクロホンアレイ間の音の到達時間差 τ を生成する。各マイクロホンアレイの定位分解能は1[deg]とし、 $\mathcal{N}(0, 4)$ [deg]の誤差を与えた。観測時間誤差は $\mathcal{N}(0, 10^{-6})$ [s]とした。サンプリング周波数は16[kHz]とした。

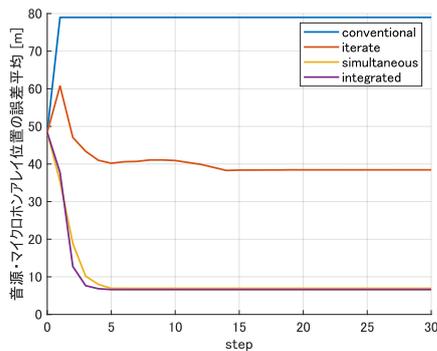


Fig. 2 音源・マイクロホンアレイ定位の誤差

Fig. 2は各最適化手法について、定位の誤差 E の推移を示したものである。

$$E = \sum_{i=1}^M \|\hat{\mathbf{S}}_i - \mathbf{S}_i\|_2 + \sum_{j=1}^N \|\hat{\mathbf{A}}_j - \mathbf{A}_j\|_2 \quad (5)$$

ただし、 \mathbf{S}, \mathbf{A} は音源およびマイクロホンアレイの座標の真値である。最適化計算において一定の評価回数ごとに1ステップとして、100通りの配置に対する誤差の

平均を記録した。従来法は1ステップ目で局所最適解に収束し、繰り返し推定は従来法と比較して誤差は小さいものの、これも局所最適解に陥っている。一方で目的関数を連立・統合した手法はどちらも7[m]程度の誤差となった。これは音源方向および観測時間差に加えた誤差によって生じる値の範囲に収まっており、正しく推定できていると考える。また、収束までにかかるステップ数が4, 5ステップであり、繰り返し推定よりも10ステップほど少ないことから、音源方向・観測時間差を同時に評価する手法は、今回の条件において計算時間が短い点でも優れている。

しかし、シミュレーションの条件では音源方向と発音時刻が高精度で計測できると仮定している。実環境における実験では壁からの反響などによって音響信号処理の精度が落ちるため、今後の課題として実際の観測を用いた推定精度の評価が挙げられる。

4 おわりに

本稿は、複数のマイクロホンアレイによる複数の音源からの音響信号の観測を用いて、音源位置およびマイクロホンアレイの位置・回転・時間オフセットを推定する問題の最適化手法について述べた。解の探索において簡単なアルゴリズムを用いる場合、マイクロホンアレイに対する音源方向とマイクロホンアレイ間の音響信号の到達時間差を同時に参照する手法は、独立して解く手法よりも精度の高い推定が行えるとわかった。本稿では入力データの誤差が小さいと仮定して性能を評価したため、今後はマイクロホンアレイの観測によって生じるノイズを考慮した推定を行い、その精度を評価することが課題となる。

謝辞 本研究はJSPS科研費JP19K12017, JP19KK0260およびJP20H00475の助成を受けた。

参考文献

- [1] 浅野 太., “音のアレイ信号処理”, コロナ社, 音響テクノロジーシリーズ/日本音響学会編, 2011.
- [2] 山田 ほか, “マイクロホンアレイ搭載ドローンによる音源方向尤度統合に基づく音源追跡”, *JSAI Technical Report, SIG-Challenge-057-02*, 2020.
- [3] D. Gabriel, et al., “2D sound source position estimation using microphone arrays and its application to a VR-based bird song analysis system”, *Advanced Robotics* 33(7-8), pp.403-414, 2019.
- [4] A. Plinge and G. A. Fink, “Geometry calibration of multiple microphone arrays in highly reverberant environments.”, *2014 14th IWAENC*, 243-247, 2014.
- [5] A. Plinge, G. A. Fink and S. Gannot, “Passive online geometry calibration of acoustic sensor networks”, *IEEE Signal Processing Letters* 24(3), pp.324-328, 2017.
- [6] S. Woźniak and K. Kowalczyk, “Passive joint localization and synchronization of distributed microphone arrays.”, *IEEE Signal Processing Letters*, 26(2), pp.292-296, 2018.