

波源拘束差分方程式に基づく音響信号の確率モデル化と 複数音源定位アルゴリズム*

☆鈴木 惇 (東大院・情報理工), 亀岡弘和 (東大院・情報理工 NTT CS 研)

1 はじめに

波源定位は, レーダやソナーといった幅広い応用を有する。特に, 小さいアレイで, 移動する波源を瞬時に定位し追跡可能にすることは重要課題である。波源定位問題に対する従来法としては, Multiple Signal Classification (MUSIC) 法 [1], Generalized Cross-Correlation methods with Phase Transform (GCC-PHAT) 法 [2], 波源拘束偏微分方程式に基づく手法 [3, 4, 5] などがある。

MUSIC 法や GCC-PHAT 法は, センサアレイの受信信号間の自己相関関数や相互相関関数といったセンサ間の到来時間差に由来する統計量に基づく手法であるため, 音源を高い精度で定位するためには大きいアレイサイズと十分長い観測時間幅を必要とする。一方, 波源拘束偏微分方程式に基づく手法は, 各時刻ごとに成立する音響信号の時空間偏微分方程式を元に音源定位を行うもので, 理論的には瞬時の小領域観測のみで波源定位を行うことが可能だが, 単一波源に対して成立する方程式を基にしているため, 複数の音源を同時に定位することはできない。また, 雑音が存在する場合など, 観測音響信号が偏微分方程式から逸脱する場合に脆弱であるという欠点を有する。

そこで本研究は, 小領域・瞬時観測による複数音源の波源定位を可能にする手法の提案を目的として, 波源拘束偏微分方程式の周波数領域表現を基にした音響信号およびその空間差分の確率分布を構築することで複数の音源と加法雑音が存在する場合の混合信号およびその空間差分の確率分布を記述し, それに基づき複数の音源の波源定位を行うための Expectation-Maximization (EM) アルゴリズムを導出する。また, シミュレーションで提案手法の有用性を検証する。

2 手法

2.1 波源拘束偏微分方程式

観測点の基準となる位置ベクトルを $\mathbf{r} = (x, y, z)^T \in \mathbb{R}^3$ とし, 単一波源の位置ベクトルを $\mathbf{r}_0 = (x_0, y_0, z_0)^T \in \mathbb{R}^3$ とする。波源の信号を $g(t)$, 音速を c とし, 単一点波源からの球面波伝播を仮定すると観測点における観測値は

$$f(\mathbf{r}, t) = \frac{1}{|\mathbf{r} - \mathbf{r}_0|} g\left(t - \frac{|\mathbf{r} - \mathbf{r}_0|}{c}\right) \quad (1)$$

と表される。観測点から波源方向へ向かう単位ベクトルを \mathbf{n} とすると, であるため, $f(\mathbf{r}, t)$ の空間微分は,

$$\begin{aligned} \nabla f(\mathbf{r}, t) &= \frac{1}{|\mathbf{r} - \mathbf{r}_0|^2} g\left(t - \frac{|\mathbf{r} - \mathbf{r}_0|}{c}\right) \mathbf{n} \\ &+ \frac{1}{c|\mathbf{r} - \mathbf{r}_0|} \dot{g}\left(t - \frac{|\mathbf{r} - \mathbf{r}_0|}{c}\right) \mathbf{n} \end{aligned} \quad (2)$$

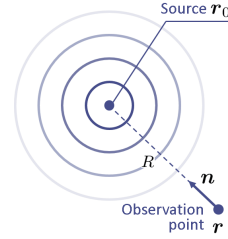


Fig. 1 点音源 \mathbf{r}_0 から観測点 \mathbf{r} へ到来する球面波。

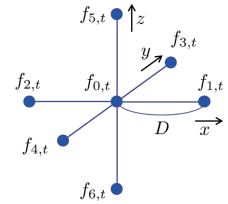


Fig. 2 観測信号 f の空間微分を取得するためのアレイ幾何の例。例えば x 方向の f の空間微分は $f_{1,t} - f_{2,t}$ で近似できる。

となる。また, $f(\mathbf{r}, t)$ の時間微分は

$$\dot{f}(\mathbf{r}, t) = \frac{1}{|\mathbf{r} - \mathbf{r}_0|} \dot{g}\left(t - \frac{|\mathbf{r} - \mathbf{r}_0|}{c}\right) \quad (3)$$

となるので, g を消去して,

$$\nabla f(\mathbf{r}, t) = \left\{ \frac{1}{R} f(\mathbf{r}, t) + \frac{1}{c} \dot{f}(\mathbf{r}, t) \right\} \mathbf{n} \quad (4)$$

のように, 観測信号とその時間・空間微分のみを含む方程式を立てられる。ただし, 距離 $R = |\mathbf{r} - \mathbf{r}_0|$ である。この式を波源拘束式と呼ぶ [3, 4, 5]。以上のように波源拘束式は, 任意の音源信号波形で成り立つ, 音源の位置と空間の場の一意な関係を記述する偏微分方程式である。

2.2 波源拘束式に基づく音響信号の確率モデル化

観測信号の空間微分を空間差分で近似する場合を考える。簡単のため, 以下では図 2 のようなマイクロホンアレイの配置を仮定する。このとき 7 本のマイクロホンを用いて各時刻 t_l で, 基準点における信号 $f_{0,l}$ およびその各方向の空間差分 $f_{x,l} = \frac{f_{1,l} - f_{2,l}}{2D}$, $f_{y,l} = \frac{f_{3,l} - f_{4,l}}{2D}$, $f_{z,l} = \frac{f_{5,l} - f_{6,l}}{2D}$ を得ることができ。ただし, l は離散時刻のインデックスを表す。基準点における観測信号の時間微分を時間差分で近似することとする, 式 (4) は

$$f_{\xi,l} = \left(\frac{1}{R} f_{0,l} + \frac{1}{c} \frac{f_{0,l} - f_{0,l-1}}{T} \right) \mathbf{n}_{\xi} \quad (5)$$

*Generative modeling of acoustic signals for multiple sound source localization based on spatio-temporal partial difference equation. by SUZUKI, Atsushi (Graduate School of Information Science and Technology, The University of Tokyo), KAMEOKA, Hirokazu (Graduate School of Information Science and Technology, The University of Tokyo NTT Communication Science Laboratories, NTT Corporation)

と表せる。ただし、 $\xi = x, y, z$ で、 n_x, n_y, n_z はそれぞれ \mathbf{n} の x, y, z 方向の成分、 T はサンプリング周期である。式 (5) の左辺を右辺に移項すると

$$-f_{\xi,l} + \left(\frac{1}{R} f_{0,l} + \frac{1}{c} \frac{f_{0,l} - f_{0,l-1}}{T} \right) n_{\xi} = 0 \quad (6)$$

が得られる。ここで、 $f_{0,l}, f_{x,l}, f_{y,l}, f_{z,l}$ を窓関数で窓掛けして取得された信号とする。切り出し区間の両端点の影響を無視すると、式 (6) は周波数領域で

$$-F_{\xi,m} + \left(\frac{1}{R} + \frac{1 - e^{-j\frac{2\pi m}{L}}}{cT} \right) n_{\xi} F_{0,m} = 0 \quad (7)$$

と表される。ただし、 $F_{0,m}, F_{x,m}, F_{y,m}, F_{z,m}$ は $f_{0,m}, f_{x,m}, f_{y,m}, f_{z,m}$ の離散 Fourier 変換であり、 m は離散周波数インデックスである。

式 (7) の右辺は雑音の存在や差分近似に伴う誤差により実際には必ずしも厳密に 0 にはならない。そこで、式 (6) の右辺を

$$-F_{\xi,m} + \left(\frac{1}{R} + \frac{1 - e^{-j\frac{2\pi m}{L}}}{cT} \right) n_{\xi} F_{0,m} = \epsilon_{\xi,m} \quad (8)$$

のように誤差変数 $\epsilon_{x,m}, \epsilon_{y,m}, \epsilon_{z,m}$ に置き換え、これらを平均が 0 で互いに独立な正規確率変数（複素正規分布に従う確率変数）

$$\epsilon_{\xi,m} \sim \mathcal{N}_{\mathbb{C}}(\epsilon_{\xi,m}; 0, \sigma_{\xi,m}^2) \quad (9)$$

と仮定する。また、観測点における観測信号の各周波数成分を平均が 0、分散が $\sigma_{0,m}^2$ の正規確率変数とする。これは、

$$F_{0,m} = \epsilon_{0,m}, \quad \epsilon_{0,m} \sim \mathcal{N}_{\mathbb{C}}(\epsilon_{0,m}; 0, \sigma_{0,m}^2) \quad (10)$$

と仮定することに相当する。ここで、

$$\mathbf{f}_m = [F_{x,m} \quad F_{y,m} \quad F_{z,m} \quad F_{0,m}]^T \in \mathbb{C}^4, \quad (11)$$

$$\boldsymbol{\epsilon}_m = [\epsilon_{x,m} \quad \epsilon_{y,m} \quad \epsilon_{z,m} \quad \epsilon_{0,m}]^T \in \mathbb{C}^4 \quad (12)$$

とし、これらを縦に連結して

$$\mathbf{f} = \begin{bmatrix} \mathbf{f}_0^T & \cdots & \mathbf{f}_{L/2}^T \end{bmatrix}^T \in \mathbb{C}^{2L+4} \quad (13)$$

$$\boldsymbol{\epsilon} = \begin{bmatrix} \boldsymbol{\epsilon}_0^T & \cdots & \boldsymbol{\epsilon}_{L/2}^T \end{bmatrix}^T \in \mathbb{C}^{2L+4} \quad (14)$$

と表記すると、式 (8) は

$$\mathbf{A}(\theta) \mathbf{f} = \boldsymbol{\epsilon} \quad (15)$$

の形で書ける。ただし、 $\theta = \{R, \mathbf{n}\}$ であり、 $\mathbf{A}(\theta)$ は

$$\mathbf{A}(\theta) = \text{blockdiag} \{ \mathbf{C}_0(\theta), \dots, \mathbf{C}_{L/2}(\theta) \}$$

$$\mathbf{C}_m(\theta) = \begin{bmatrix} -1 & & & \left(\frac{1}{R} + \frac{1 - e^{-j\frac{2\pi m}{L}}}{cT} \right) n_x \\ & -1 & & \left(\frac{1}{R} + \frac{1 - e^{-j\frac{2\pi m}{L}}}{cT} \right) n_y \\ & & -1 & \left(\frac{1}{R} + \frac{1 - e^{-j\frac{2\pi m}{L}}}{cT} \right) n_z \\ & & & 1 \end{bmatrix} \in \mathbb{C}^{4 \times 4} \quad (16)$$

で与えられる。ただし $\text{blockdiag}\{\}$ は括弧内の行列を対角に並べたブロック対角行列。式 (9), (10) より、 $\boldsymbol{\epsilon}$ は平均が $\mathbf{0}$ 、分散共分散行列が $\boldsymbol{\Sigma}$

$$\boldsymbol{\Sigma} = \text{blockdiag} \{ \boldsymbol{\Sigma}_0, \dots, \boldsymbol{\Sigma}_{L/2} \}$$

$$\boldsymbol{\Sigma}_m = \text{diag} \{ \sigma_{x,m}^2, \sigma_{y,m}^2, \sigma_{z,m}^2, \sigma_{0,m}^2 \} \quad (17)$$

の複素正規分布

$$\boldsymbol{\epsilon} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{\epsilon}; \mathbf{0}, \boldsymbol{\Sigma}) \quad (18)$$

に従う。 $\det \mathbf{A}(\theta) = (-1)^{L/2+1} \neq 0$ ($\mathbf{A}(\theta)$ は正則) であるので、 \mathbf{f} は

$$\mathbf{f} = \mathbf{A}(\theta)^{-1} \boldsymbol{\epsilon} \quad (19)$$

と表され、式 (18) より、

$$\mathbf{f} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{f}; \mathbf{0}, \mathbf{A}(\theta)^{-1} \boldsymbol{\Sigma} \mathbf{A}(\theta)^{-H}) \quad (20)$$

が言える。従って、観測信号およびその空間差分が与えられた下での最尤音源位置 $\hat{\theta}$ は

$$\hat{\theta} = \underset{\theta}{\text{argmax}} \log p(\mathbf{f}|\theta) \quad (21)$$

$$p(\mathbf{f}|\theta) = \mathcal{N}_{\mathbb{C}}(\mathbf{f}; \mathbf{0}, \mathbf{A}(\theta)^{-1} \boldsymbol{\Sigma} \mathbf{A}(\theta)^{-H}) \quad (22)$$

により得られる。

2.3 複数音源の定位アルゴリズム

以上の \mathbf{f} の確率モデル化により、音源が複数個存在する場合、および雑音が存在する場合の観測信号の確率分布が導ける。音源インデックスを k とし、音源 k に由来する観測信号の成分、音源位置パラメータをそれぞれ $\mathbf{f}^{(k)}, \theta^{(k)}$ とする。また、 $\mathbf{f}^{(k)}$ の周波数 m の成分エネルギーを $\sigma_{0,m}^{(k)2}$ とする。式 (23) より、

$$\mathbf{f}^{(k)} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{f}^{(k)}; \mathbf{0}, \mathbf{A}(\theta^{(k)})^{-1} \boldsymbol{\Sigma}^{(k)} \mathbf{A}(\theta^{(k)})^{-H}) \quad (23)$$

となる。また、加法雑音を \mathbf{v} とし、観測信号を $\mathbf{y} = \sum_k \mathbf{f}^{(k)} + \mathbf{v}$ とする。 $\mathbf{f}^{(1)}, \dots, \mathbf{f}^{(K)}, \mathbf{v}$ が互いに独立であれば、観測信号 \mathbf{y} は $\mathbf{y} \sim p(\mathbf{y}|\theta)$

$$p(\mathbf{y}|\theta) = \mathcal{N}_{\mathbb{C}}(\mathbf{y}; \mathbf{0}, \sum_k \mathbf{A}(\theta^{(k)})^{-1} \boldsymbol{\Sigma}^{(k)} \mathbf{A}(\theta^{(k)})^{-H} + \boldsymbol{\Gamma})$$

$$= \prod_m \mathcal{N}_{\mathbb{C}}(\mathbf{y}_m; \mathbf{0}, \sum_k \mathbf{C}_m(\theta^{(k)})^{-1} \boldsymbol{\Sigma}_m^{(k)} \mathbf{C}_m(\theta^{(k)})^{-H} + \boldsymbol{\Gamma}_m) \quad (24)$$

に従う。ただし、 $\boldsymbol{\Gamma}$ は \mathbf{v} の分散共分散行列である。以上より、複数の音源と雑音が存在する場合の各音源の最尤音源位置 $\hat{\theta} = \{\hat{\theta}^{(1)}, \dots, \hat{\theta}^{(K)}\}$ は、観測信号 $\mathbf{y} = (\mathbf{y}_0^T, \dots, \mathbf{y}_{L/2}^T)^T$ が与えられた下で

$$\hat{\theta} = \underset{\theta}{\text{argmax}} p(\mathbf{y}|\theta) \quad (25)$$

を解くことにより得られる。

\mathbf{y} を不完全データ、 $\mathbf{x} = (\mathbf{f}^{(1)T}, \dots, \mathbf{f}^{(K)T}, \mathbf{v}^T)^T$ を完全データと見なすと、以上の最尤推定問題に対し Expectation-Maximization (EM) アルゴリズムを適用できる。完全データ対数尤度 $\log p(\mathbf{x}|\theta)$ は

$$\log p(\mathbf{x}|\theta) = -\log |\det \boldsymbol{\Lambda}| - \mathbf{x}^H \boldsymbol{\Lambda}^{-1} \mathbf{x} \quad (26)$$

$$\Lambda = \begin{bmatrix} A(\theta^{(1)})^{-1} \Sigma^{(1)} A(\theta^{(1)})^{-H} & & & & O \\ & \ddots & & & \\ & & A(\theta^{(K)})^{-1} \Sigma^{(K)} A(\theta^{(K)})^{-H} & & \\ O & & & & \Gamma \end{bmatrix} \quad (27)$$

与えられるので、 \mathbf{y} が与えられた下での $\log p(\mathbf{x}|\theta)$ の \mathbf{x} に関する条件付き期待値 (Q 関数) は

$$Q(\theta, \theta') = \mathbb{E}[\log p(\mathbf{x}|\theta)|\mathbf{y}, \theta'] \quad (28)$$

$$= -\log |\det \pi \Lambda| - \text{tr}(\Lambda^{-1} \mathbb{E}[\mathbf{x}\mathbf{x}^H|\mathbf{y}, \theta']) \quad (29)$$

$$= \frac{\theta}{m} - \sum_k \log |\Sigma^{(k)}| - \log |\Gamma| - \text{tr}(\Gamma^{-1} \mathbb{E}[\mathbf{v}\mathbf{v}^H|\mathbf{y}, \theta']) - \sum_k \text{tr}(\mathbf{A}(\theta^{(k)})^H \Sigma^{(k)-1} \mathbf{A}(\theta^{(k)}) \mathbb{E}[\mathbf{f}^{(k)} \mathbf{f}^{(k)H}|\mathbf{y}, \theta'])$$

与えられる。ただし、 $\frac{\theta}{m}$ は x に関する項のみについての等号を意味する。この関数が増大するように θ を更新するステップ (M ステップ) と、更新した θ を θ' に代入し、 $\mathbb{E}[\mathbf{f}^{(k)} \mathbf{f}^{(k)H}|\mathbf{y}, \theta']$ と $\mathbb{E}[\mathbf{v}\mathbf{v}^H|\mathbf{y}, \theta']$ を計算するステップ (E ステップ) を繰り返すことで $p(\mathbf{y}|\theta)$ を局所的に最大にする θ を求められる。

完全データ \mathbf{x} と不完全データ \mathbf{y} の関係は

$$\mathbf{y} = \mathbf{H}\mathbf{x}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{I} & \cdots & \mathbf{I} & \mathbf{I} \end{bmatrix} \quad (30)$$

と書けるので、 $\mathbb{E}[\mathbf{x}|\mathbf{y}, \theta']$ と $\mathbb{E}[\mathbf{x}\mathbf{x}^H|\mathbf{y}, \theta']$ はそれぞれ

$$\begin{aligned} \mathbb{E}[\mathbf{x}|\mathbf{y}, \theta'] &= \Lambda \mathbf{H}^T (\mathbf{H} \Lambda \mathbf{H}^T)^{-1} \mathbf{y} \\ \mathbb{E}[\mathbf{x}\mathbf{x}^H|\mathbf{y}, \theta'] &= \Lambda - \Lambda \mathbf{H}^T (\mathbf{H} \Lambda \mathbf{H}^T)^{-1} \mathbf{H} \Lambda \\ &\quad + \mathbb{E}[\mathbf{x}|\mathbf{y}, \theta'] \mathbb{E}[\mathbf{x}|\mathbf{y}, \theta']^H \\ &= \begin{bmatrix} \Phi^{(1)} & & & * \\ & \ddots & & \\ & & \Phi^{(K)} & \\ * & & & \Phi^{(K+1)} \end{bmatrix} \end{aligned} \quad (31)$$

与えられる。以上より、以下のアルゴリズムを得る。

1. θ を初期設定する。
2. (E ステップ) θ を θ' に代入し、式 (31) により $\Phi^{(1)}, \dots, \Phi^{(K)}, \Phi^{(K+1)}$ を計算する。
3. (M ステップ) 下式により θ を更新する。

$$\theta \leftarrow \underset{\theta}{\text{argmax}} Q(\theta, \theta') \quad (32)$$

$$\begin{aligned} Q(\theta, \theta') &= -\sum_k \log |\Sigma^{(k)}| - \log |\Gamma| \\ &\quad - \sum_k \text{tr}(\mathbf{A}(\theta^{(k)})^H \Sigma^{(k)-1} \mathbf{A}(\theta^{(k)}) \Phi^{(k)}) \\ &\quad - \text{tr}(\Gamma^{-1} \Phi^{(K+1)}) \end{aligned} \quad (33)$$

2.4 M ステップ更新式

$\mathbf{A}(\theta^{(k)})$ は $\mathbf{C}_1(\theta^{(k)}), \dots, \mathbf{C}_L(\theta^{(k)})$ を対角に並べたブロック対角行列なので、

$$\begin{aligned} &\text{tr}(\mathbf{A}(\theta^{(k)})^H \Sigma^{(k)-1} \mathbf{A}(\theta^{(k)}) \Phi^{(k)}) \\ &= \sum_m \text{tr}(\mathbf{C}_m(\theta^{(k)})^H \Sigma_m^{(k)-1} \mathbf{C}_m(\theta^{(k)}) \Phi_m^{(k)}) \end{aligned} \quad (34)$$

と書かれる。ただし、

$$\Sigma_m^{(k)} = \text{diag} \left\{ \sigma_{x,m}^{(k)2}, \sigma_{y,m}^{(k)2}, \sigma_{z,m}^{(k)2}, \sigma_{0,m}^{(k)2} \right\} \quad (35)$$

であり、 $\Phi_m^{(k)}$ は $\Phi^{(k)}$ の 4×4 のブロック対角成分。

M ステップでは $Q(\theta, \theta')$ ができるだけ大きくなるように $\mathbf{n}^{(k)}, R^{(k)}, \sigma_{0,m}^{(k)2}, \Gamma$ を更新する。 $Q(\theta, \theta')$ を最大にする $\mathbf{n}^{(k)}, R^{(k)}, \sigma_{0,m}^{(k)2}, \Gamma$ の同時最適解を解析的に求めることは難しいが、座標勾配法によりそれぞれの変数に関して $Q(\theta, \theta')$ が最大となるように反復更新することで $p(\mathbf{y}|\theta)$ を局所最大化することができる (EM アルゴリズムでは、M ステップで補助関数が単調に増大することが保証されていれば収束性は保証される)。以下に M ステップの更新方法を示す。

1) $\mathbf{n}^{(k)}$ の更新式

$\mathbf{n}^{(k)}$ は単位ベクトルなので、 $\|\mathbf{n}^{(k)}\|_2^2 = 1$ の下で $\text{tr}(\mathbf{A}(\theta^{(k)})^H \Sigma^{(k)-1} \mathbf{A}(\theta^{(k)}) \Phi^{(k)})$ ができるだけ小さくなるように $\mathbf{n}^{(k)}$ を更新する。この制約つき最適化問題は、目的関数

$$\begin{aligned} \mathcal{L}(\mathbf{n}^{(k)}) &= \text{tr}(\mathbf{A}(\theta^{(k)})^H \Sigma^{(k)-1} \mathbf{A}(\theta^{(k)}) \Phi^{(k)}) \\ &= \sum_m \text{tr}(\mathbf{C}_m(\theta^{(k)})^H \Sigma_m^{(k)-1} \mathbf{C}_m(\theta^{(k)}) \Phi_m^{(k)}) \end{aligned} \quad (36)$$

を $\mathbf{n}^{(k)T} \mathbf{n}^{(k)} - 1 = 0$ の条件のもと最小化する問題と等価である。この問題は二次制約付き二次最適化問題であり、Lagrange 未定乗数法などで解ける。

2) 音源距離 $R^{(k)}$ の更新式

$\rho^{(k)} = \frac{1}{R^{(k)}}$ とする。上記と同様、 $\mathbf{C}_m(\theta^{(k)})$ は

$$\mathbf{C}_m(\theta^{(k)}) = \rho^{(k)} \mathbf{V}^{(k)} + \mathbf{P}_m^{(k)} \quad (38)$$

$$\mathbf{V}^{(k)} = \begin{bmatrix} 0 & 0 & 0 & n_x^{(k)} \\ 0 & 0 & 0 & n_y^{(k)} \\ 0 & 0 & 0 & n_z^{(k)} \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (39)$$

$$\mathbf{P}_m^{(k)} = \begin{bmatrix} -1 & & & \frac{1-e^{-j\frac{2\pi m}{L}}}{cT} n_x^{(k)} \\ & -1 & & \frac{1-e^{-j\frac{2\pi m}{L}}}{cT} n_y^{(k)} \\ & & -1 & \frac{1-e^{-j\frac{2\pi m}{L}}}{cT} n_z^{(k)} \\ & & & 1 \end{bmatrix} \quad (40)$$

と、 $\rho^{(k)}$ に依存する項としない項に分解できるので、 $Q(\theta, \theta')$ の $\rho^{(k)}$ に関する偏微分を 0 と置くことで、

$$\rho^{(k)} = -\frac{\sum_m \text{tr}(\Re(\mathbf{V}^{(k)T} \Sigma_m^{(k)-1} \mathbf{P}_m^{(k)} \Phi_m^{(k)}))}{\sum_m \text{tr}(\mathbf{V}^{(k)T} \Sigma_m^{(k)-1} \mathbf{V}^{(k)} \Phi_m^{(k)})} \quad (41)$$

を得る。

3) $\sigma_{0,m}^{(k)2}$ の更新式

上記と同様、 $\Sigma_m^{(k)-1}$ を

$$\Sigma_m^{(k)-1} = \mathbf{G}_m^{(k)} + \frac{1}{\sigma_{0,m}^{(k)2}} \mathbf{E}_{4,4} \quad (42)$$

と、 $\sigma_{0,m}^{(k)2}$ に依存する項としない項に分解できるので、 $Q(\theta, \theta')$ の $\sigma_{0,m}^{(k)2}$ に関する偏微分を 0 と置くことで、

$$\sigma_{0,m}^{(k)2} = [\Phi_m^{(k)}]_{4,4} \quad (43)$$

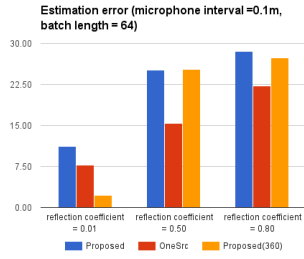


Fig. 3 フレーム幅が $L = 64$ の場合の定位精度。

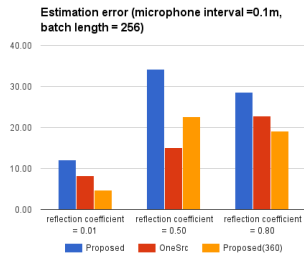


Fig. 4 フレーム幅が $L = 256$ の場合の定位精度。

を得る。

4) 雑音分散共分散行列 Γ の更新式 雑音の分散共分散行列を

$$\Gamma = \text{blockdiag} \left\{ \nu_0^2 \mathbf{W}_0, \dots, \nu_{L/2}^2 \mathbf{W}_{L/2} \right\} \quad (44)$$

のように、正規化分散共分散行列モデル \mathbf{W}_m と周波数 k の成分エネルギー ν_m^2 の積で表し、 ν_m^2 を推定すべき変数とする。ここで \mathbf{W}_m は空間無相関モデルや、マイクの 2 点間の空間相関係数が

$$\eta_d(\omega) = \frac{\sin \frac{\omega d}{c}}{\frac{\omega d}{c}} \quad (45)$$

で与えられる拡散音場モデルなどから導かれる定数行列。 $Q(\theta, \theta')$ の ν^2 に関する偏微分を 0 と置くことで、

$$\nu_m^2 = \frac{1}{4} \text{tr}(\mathbf{W}_m^{-1} \Phi_m^{(K+1)}) \quad (46)$$

を得る。

3 数値実験

提案手法の性能を検証するため、残響環境下の 1 音源定位の数値シミュレーションを行った。今回、 x 方向、 y 方向の 2 次元モデルを使用した。部屋サイズは $6\text{m} \times 10\text{m} \times 8\text{m}$ とし、中心に 0.1m 間隔でマイクを 5 つ配置した。壁面の反射係数は 0.01, 0.5, 0.8 (残響の影響の大きさに相当) の 3 条件とした。サンプリング周波数は 16000, フレーム幅は $L = 64, 256$ とした。音源の配置は図 5 の通り。それぞれの条件下で 40 種類の音源を発させ、到来角度を推定した際の誤差の二乗和平均の平方根 (deg) を評価した。通常の EM アルゴリズムによる手法が図中の “Proposed” である。単一音源のみの存在を仮定した尤度関数 (式 (22)) を最大にする方法 (従来法に相当・図中の “OneSrc”) と比較している。図 3, 4 に実験結果を示す。

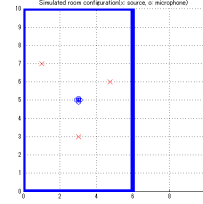


Fig. 5 音源位置とマイク位置。赤バツ印が音源位置 (別々に実験)。青マル印がマイク位置。

EM アルゴリズムは、望ましくない局所解に陥ることがあり、それを避けるために 360 点のランダムな方向から最尤なものを初期解する手法 (図中の “Proposed360”) についても実験を行った。初期解を吟味して与えた場合、提案手法は比較手法と比べ、高精度な定位が可能ながある場合があり、MUSIC 法や GCC-PHAT が使えない短時間測定からの定位問題で有用であることが分かる。

4 結論

本研究は、波源拘束偏微分方程式の確率モデル化に基づき小領域・瞬時観測による複数音源の定位を可能にする手法を提案し、数値実験で有用性を示した。

謝辞 本研究は JSPS 科研費 26730100 の助成を受けて行われた。

参考文献

- [1] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” IEEE Transactions on Antennas and Propagation, vol. 34, no. 3, pp. 276–280, 1986.
- [2] C. H. Knapp et al., “The generalized correlation method for estimation of time delay,” IEEE Trans. ASSP, vol. 24, no. 4, pp. 320–327, 1976.
- [3] 藤田悠哉, 小野順貴, 安藤繁, “有限時間窓と離散フーリエ変換の利用を可能にする音源定位の高速厳密解法とその実験” 日本音響学会 2006 年秋季研究発表会講演論文集, 3-1-3, pp. 483–484, Sep. 2006.
- [4] S. Ando, N. Ono, T. Nara, “Direct algebraic method for sound source localization with finest resolution both in time and frequency,” in Proc. ICSV14, Jul. 2007.
- [5] 小山翔一, 栗原徹, 安藤繁, “偏微分方程式の空間荷重積分による瞬時音源定位,” 日本音響学会 2008 年秋季研究発表会講演論文集, 2-8-20, pp. 679–682, Sep. 2008.