

分離信号間のエネルギー差に基づいて音源方向を決定する劣決定音源分離*

◎ 皆川朋樹, 矢田部浩平 (農工大)

1 はじめに

劣決定な音源分離に対して、事前に定義した空間モデルに基づく凸最適化問題として定式化する手法が提案されており、音源到来方向 (DOA) と音源を同時に推定できる [1]。この手法では平面波モデルを前提とし、グリッド状の方位角に対応するステアリングベクトルをまとめた空間辞書を用いる。これを用いて信号を方向別の仮想音源に分離し、エネルギーが集中する方位角を推定 DOA、対応する仮想音源を推定音源とする。しかし、高い推定精度を得るためにはグリッドを細かくする必要があり、それに伴って辞書行列が大きくなることから、計算コストが増大する課題がある。これに対し我々は、計算コスト削減のために限定的な方位角のみを用いる手法を提案した [2]。同手法では、分離信号間のエネルギーの大小関係に基づいて方位角を逐次更新することで DOA と音源を推定する。本稿では、処理のさらなる効率化を目的とし、分離信号間のエネルギー差を利用して更新ステップ幅を適応的に決定する手法を提案する。

2 空間辞書を用いた音源分離

N 個の音源 $\mathbf{s}(t, f) = [s_1(t, f), \dots, s_N(t, f)]^T \in \mathbb{C}^N$ が、位置が既知の M 個のマイクロホン $\mathbf{x}(t, f) = [x_1(t, f), \dots, x_M(t, f)]^T \in \mathbb{C}^M$ で観測された状況を考える。ただし、 t と f はそれぞれ時間と周波数のインデックス、 $(\cdot)^T$ は転置である。ここで、マイクロホン位置と K 個の方位角 $\boldsymbol{\theta} = [\theta_1, \dots, \theta_K]^T \in \mathbb{R}^K$ ($K > N$) に対応するステアリングベクトルをまとめた空間辞書 $\mathbf{A}(f) \in \mathbb{C}^{M \times K}$ 、 $\boldsymbol{\theta}$ に対応する仮想音源 $\mathbf{y}(t, f) = [y_1(t, f), \dots, y_K(t, f)]^T \in \mathbb{C}^K$ を与えると、音源分離問題を以下の凸最適化問題で定式化できる [1]。

$$\underset{\mathbf{y}}{\text{Minimize}} \quad \|\mathbf{x} - \mathbf{A}\mathbf{y}\|_2^2 + \mu \|\mathbf{y}\|_{2,1} + \nu \|\mathbf{y}\|_{1,2} \quad (1)$$

ただし、 $\mu, \nu > 0$ は正則化パラメータである。この問題の解には θ_k に対する分離信号が含まれているため、方位角数 K が音源数 N より十分大きい場合は、分離信号のエネルギーが大きな θ_k を N 個選択することで DOA を推定し、 $s_n(t, f) \approx y_k(t, f)$ として仮想音源から音源を推定する。一方、高い推定精度を得るためにはグリッド数 K を大きくする必要があり、計算コストとのトレードオフが生じる。これに対

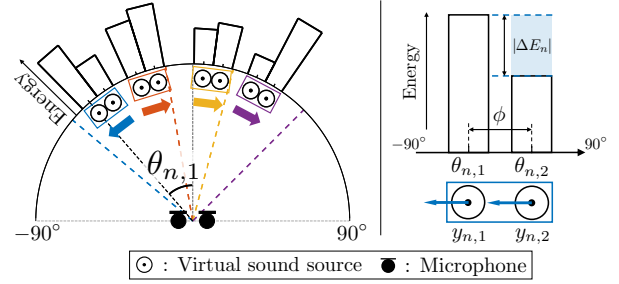


図-1 全方位角の更新の様子 (左) と、1 組のペアの更新の様子 (右)。仮想音源のエネルギーを棒グラフで、音源方向とそれに対応する方位角の更新方向をそれぞれ色付けされた破線と矢印で示した。

し [2] では、各音源に対し 2 つの方位角を組とする方位角ペア $(\theta_{n,1}, \theta_{n,2})$ を用意し ($K = 2N$)、これらを逐次更新することで計算コストを抑制する。ここで、ペア内における 2 要素の角度間隔を ϕ とし、式 (1) から得られる各仮想音源の分離信号エネルギーを $E_{n,l} = \sum_{t,f} |y_{n,l}(t, f)|^2$ ($l = 1, 2$) と定義する。同手法では、図-1 (左) のように $E_{n,1}$ と $E_{n,2}$ の大小関係に基づき、音源に近い方の方位角へ各ペアを固定の更新幅で近づけることで DOA を推定する。

3 提案法

前節で述べた手法 [2] では、エネルギーの大小関係に基づき、方位角ベクトル $\boldsymbol{\theta} = [\theta_{1,1}, \theta_{1,2}, \dots, \theta_{N,2}]^T \in \mathbb{R}^{2N}$ を固定の更新幅で更新した。しかし、更新幅が小さい場合には収束までの更新回数が増大し、逆に大きい場合には DOA の推定精度が低下するという課題があった。そこで本稿では、図-1 (右) のようにエネルギー差 $\Delta E_n = E_{n,2} - E_{n,1}$ に基づき、更新幅を適応的に決定する手法を提案する。提案法における各方位角の更新幅 $\boldsymbol{\psi} = [\psi_1, \dots, \psi_N]^T$ と更新式を以下のように与える。

$$\psi_n^{[i]} = \text{sgn}(\Delta E_n^{[i]}) \cdot \min\left(\eta_n^{[i]} \frac{|\Delta E_n^{[i]}|}{\phi}, \psi_{\max}\right) \quad (2)$$

$$\theta_{n,l}^{[i+1]} = \theta_{n,l}^{[i]} + \psi_n^{[i]} \quad (3)$$

ただし、 $\text{sgn}(\cdot)$ は符号関数、 $\boldsymbol{\eta} = [\eta_1, \dots, \eta_N]^T$ ($\eta_n > 0$) はステップサイズ、 $\psi_{\max} > 0$ は推定値の急激な変動を抑制するための最大更新幅であり、 $(\cdot)^{[i]}$ は i 回更新したときの値とする。提案法では、方位角ペアにおけるエネルギー差分を更新方向の指標とし、勾配降下法に準じた方位角の更新を行う。

* Underdetermined source separation with source direction estimation based on energy differences of separated signals. By Tomoki MINAGAWA and Kohei YATABE (Tokyo University of Agriculture and Technology).

Algorithm 1 Proposed Method

Input: $\theta^{[0]}, \mathbf{x}, \eta^{[0]}, N, \phi, \psi_{\max}, \psi_{\min}, \gamma$
Output: $\hat{\mathbf{y}}, \hat{\theta}$

```

1:  $i \leftarrow 0$ 
2:  $\psi_n^{[0]} \leftarrow \infty$  (for all  $n$ )
3: while  $\max_n |\psi_n^{[i]}| \geq \psi_{\min}$  do
4:    $\mathbf{y}^{[i]} = \text{solveProblem}(\theta^{[i]}, \mathbf{x})$  (see Eq. (1))
5:   if  $i > 0$  then
6:      $\eta^{[i]} = \text{updateEta}(\mathbf{y}^{[i]}, \mathbf{y}^{[i-1]}, \eta^{[i-1]}, \gamma)$  (see Eq. (5))
7:   end if
8:    $\psi^{[i]} = \text{updatePsi}(\mathbf{y}^{[i]}, \eta^{[i]}, \phi, \psi_{\max})$  (see Eq. (2))
9:    $\theta^{[i+1]} = \text{updateTheta}(\theta^{[i]}, \psi^{[i]})$  (see Eq. (3))
10:   $i \leftarrow i + 1$ 
11: end while
12:  $\hat{\theta} = \text{determineDOA}(\theta^{[i]}, N)$  (see Eq. (6))
13:  $\hat{\mathbf{y}} = \text{solveProblem}(\hat{\theta}, \mathbf{x})$  (see Eq. (1))

```

式 (1) による分離と式 (3) による更新を繰り返すと、方位角がエネルギーのピークへ近づくにつれて $|\Delta E_n|$ が減少し、これに伴い $|\psi_n|$ も減少する。これを利用して以下の収束判定を行う。

$$\text{Terminate if } \max_n |\psi_n^{[i]}| < \psi_{\min} \quad (4)$$

ただし、 $\psi_{\min} > 0$ は収束判定の閾値であり、所望の DOA 推定精度に応じて設定する。また、式 (4) の収束の安定化と推定値の振動抑制を目的とし、式 (2) のステップサイズ η_n について以下の更新則を導入する。

$$\eta_n^{[i]} = \begin{cases} \gamma \eta_n^{[i-1]} & (\text{sgn}(\Delta E_n^{[i]}) \neq \text{sgn}(\Delta E_n^{[i-1]})) \\ \eta_n^{[i-1]} & (\text{otherwise}) \end{cases} \quad (5)$$

ただし、 $0 < \gamma < 1$ は減衰係数であり、更新方向の反転のたびにステップサイズを縮小させることで収束点付近での安定した収束を実現する。式 (4) を満たした際は、最終的な DOA 推定値を次式で決定する。

$$\hat{\theta}_n = \frac{\theta_{n,1}^{[i]} + \theta_{n,2}^{[i]}}{2} \quad (6)$$

最後に、各推定 DOA に対する分離音を得るため、 $K = N$ とし、推定した $\hat{\theta} = [\hat{\theta}_1, \dots, \hat{\theta}_N]^T$ に基づく空間辞書を用いて式 (1) を再度解くことで、最終的な推定音源 $s_n(t, f) \approx \hat{y}_n(t, f)$ を得る。

提案法のアルゴリズムを **Alg. 1** にまとめる。なお、方位角の収束先 $\hat{\theta}$ は初期値 $\theta^{[0]}$ に依存するため、その初期値は [2] と同様の手順に従って決定する。

4 実験

提案法の評価のため、SiSEC 2011 の dev1 liverec より全 16 パターンの 4 話者混合音声の分離実験を行った。音源数 $N = 4$ 、マイクロホン数 $M = 2$ とし、マイク間隔は 5 cm である。また、各音源の真の DOA は $-50^\circ, -10^\circ, 15^\circ, 45^\circ$ である。残響時間は 130 ms とした。また、方位角の間隔 ϕ は 1° 、最大更新幅 $\psi_{\max}$ は 15° 、収束判定の閾値 $\psi_{\min}$ は $5^\circ - 0.01^\circ$ の 9 パターン、ステップサイズの初期値 $\eta_n^{[0]}$ は 50、減衰係数 γ は 0.5

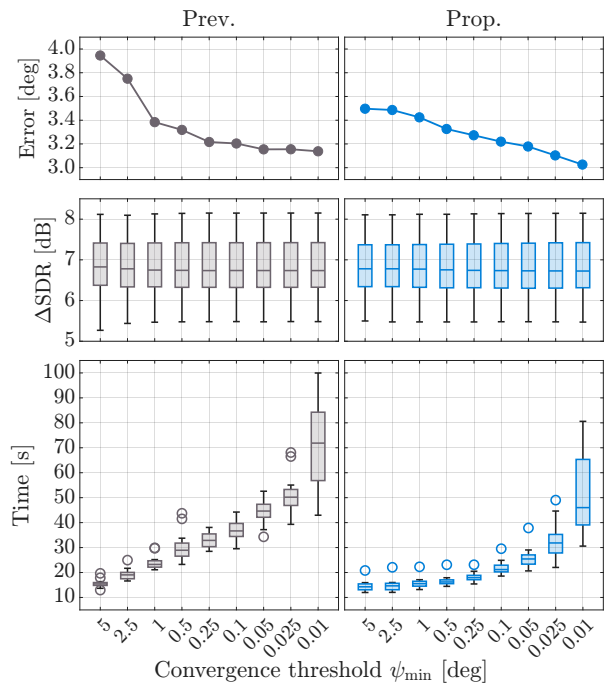


図-2 比較法と提案法における推定 DOA の平均絶対誤差 (上)、各音源の ΔSDR の平均 (中央)、処理時間 (下)。

とした。比較対象には、[2] を提案法と対照的になるよう拡張した手法を用いた。具体的には、提案法との共通の枠組みにおいて $\eta_n^{[0]} = 1$ とし、式 (2) におけるエネルギー差の項 $|\Delta E_n^{[i]}|/\phi$ を定数 $\psi_{\text{const}} = 5$ で置き換えた手法である。これにより、更新幅 ψ_n がエネルギー差 ΔE_n によらず、大小関係 $\text{sgn}(\Delta E_n)$ のみによって決定される条件を設定した。

実験結果を図-2 に示す。手法間を比較すると、DOA 推定精度において提案法は比較法よりも誤差の最大値・最小値の双方を低減させた。また、処理時間についても、提案法は $\psi_{\min}$ が小さい条件における時間の増加を抑制した。これは、更新幅を適応的に決定することで収束に必要な更新回数が効率的に削減されたためであると考えられる。

5 むすび

本稿では、音源方向を決定する際に分離信号間のエネルギー差を利用する手法を提案した。実験の結果、比較法と同等の分離性能を維持しつつ、より高い DOA 推定精度と計算コストの低減を両立した。今後は、パラメータ設定の自動化に取り組む予定である。

参考文献

- [1] T. Tachikawa, K. Yatabe, and Y. Oikawa, “Underdetermined source separation with simultaneous DOA estimation without initial value dependency,” *Int. Workshop Acoust. Signal Enhanc. (IWAENC)*, pp. 161–165 (2018).
- [2] 皆川朋樹, 矢田部浩平, “凸最適化に基づく劣決定音源分離における支配的な音源方向の選択法,” 日本音響学会講演論文集, pp. 343–344 (2025.3).
- [3] 皆川朋樹, 松本和樹, 山田宏樹, 矢田部浩平, “凸最適化に基づく劣決定 BSS に対する ADMM アルゴリズムの適用,” 日本音響学会講演論文集, pp. 195–196 (2024.9).