

空間辞書を用いた劣決定音源分離の方位角更新時のステップサイズ制御*

© 皆川朋樹, 吉野夏樹, 矢田部浩平 (農工大)

1 はじめに

劣決定な音源分離と音源到来方向 (DOA) の推定を, 空間辞書を用いた凸最適化問題として同時に行う手法が提案されている [1]. これに対し我々は, 音源ごとに 2 個の方位角をペアとして設定し, 分離信号のエネルギー差に基づいて方位角を逐次更新することで, 計算コストを削減する手法を提案した [2]. 同手法では方位角の安定収束を目指し, 更新方向が反転した際のみステップサイズを縮小し, それ以外は維持する条件分岐的な制御を行っていた. 本稿では処理の更なる効率化に向け, エネルギーの平均変化率の変動に応じてステップサイズを連続的かつ動的に拡大・縮小する手法を提案する. 音声分離実験を通じ, 本手法が推定精度を維持しながら更新回数を削減できることを示した.

2 空間辞書を用いた音源分離と方位角更新

本稿では, N 個の音源に対して $2N$ 個の方位角 $\theta = [\theta_{1,1}, \theta_{1,2}, \dots, \theta_{N,1}, \theta_{N,2}]^T \in \mathbb{R}^{2N}$ を設定し, 方位角ペア $(\theta_{n,1}, \theta_{n,2})$ を図-1 の矢印で示すように移動させながら DOA と音源信号を同時に推定する手法 [2] をベースとする. ただし, t と f は時間と周波数のインデックス, $(\cdot)^T$ は転置を表す. M 個のマイクロホンによる観測信号を $\mathbf{x}_{t,f} \in \mathbb{C}^M$, 方位角 θ に対応するステアリングベクトルを並べた空間辞書を $\mathbf{A}_f \in \mathbb{C}^{M \times 2N}$ とすると, 各方位角に対応する分離信号 $\mathbf{y}_{t,f} = [y_{t,f}^{(\theta_{1,1})}, y_{t,f}^{(\theta_{1,2})}, \dots, y_{t,f}^{(\theta_{N,1})}, y_{t,f}^{(\theta_{N,2})}]^T \in \mathbb{C}^{2N}$ は以下の凸最適化問題により得られる [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, $\mu, \nu > 0$ は解のスパース性を制御する正則化パラメータである.

本手法における処理フローを図-2 に示す. まず, 式 (1) を解くことで分離信号 $\mathbf{y}_{t,f}$ を得る. 次に, 図-1 に示す方位角の更新を, 次式により与えられるエネルギーの平均変化率 d_n を用いて行う.

$$d_n = \frac{\bar{E}^{(\theta_{n,2})} - \bar{E}^{(\theta_{n,1})}}{\theta_{n,2} - \theta_{n,1}} = \frac{\Delta \bar{E}_n}{\phi}$$

ただし, $\bar{E}^{(\theta_{n,i})} = E^{(\theta_{n,i})} / (\sum_{n',i'} (E^{(\theta_{n',i'})})^2)^{\frac{1}{2}}$ は, 各分離信号のエネルギー $E^{(\theta_{n,i})} = \sum_{t,f} |y_{t,f}^{(\theta_{n,i})}|^2$ を,

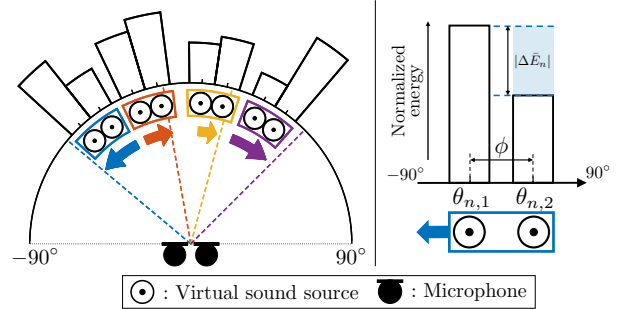


図-1 全方位角の更新の様子 (左) と, 1 組のペアの更新の様子 (右). 分離信号のエネルギーを棒グラフで, 真の DOA とそれに対応する方位角の更新方向をそれぞれ色付けされた破線と矢印で示した.

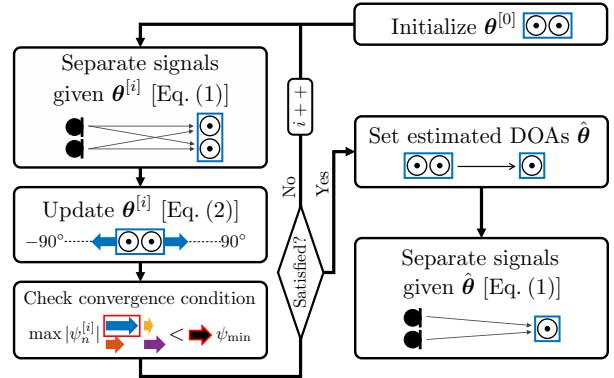


図-2 方位角更新による DOA と音源信号の推定フロー.

エネルギー全体の ℓ_2 ノルムで正規化した値である. この平均変化率 d_n に基づき, 方位角の更新式を

$$\begin{aligned} \psi_n^{[i]} &= \text{sgn}(d_n^{[i]}) \cdot \min(\eta_n^{[i]} |d_n^{[i]}|, \psi_{\max}) \\ \theta_{n,l}^{[i+1]} &= \theta_{n,l}^{[i]} + \psi_n^{[i]} \end{aligned} \quad (2)$$

と与える [2]. ただし, $\psi = [\psi_1, \dots, \psi_N]^T \in \mathbb{R}^N$ は更新量, $\eta = [\eta_1, \dots, \eta_N]^T \in \mathbb{R}_{++}^N$ はステップサイズ, $\psi_{\max} > 0$ は急激な変動を抑制するための最大更新幅, $\text{sgn}(\cdot)$ は符号関数であり, $(\cdot)^{[i]}$ は第 i 反復における値を表す. 最後に, 閾値 $\psi_{\min} > 0$ を用いた収束条件 $\max_n |\psi_n^{[i]}| < \psi_{\min}$ を満たすまで処理を反復する. 条件を満たした時点の方位角 $\theta^{[i]}$ から推定 DOA を決定し, その推定 DOA のもとで再び式 (1) を解くことで推定音源の分離信号を得る [2].

3 方位角更新のステップサイズ制御

従来法 [2] では, 式 (2) におけるステップサイズ η_n に対し, 方位角の更新方向が反転した際 ($d_n^{[i]} d_n^{[i-1]} < 0$)

*Step-size control for azimuth update in spatial dictionary-based underdetermined source separation. By Tomoki MINAGAWA, Natsuki YOSHINO and Kohei YATABE (Tokyo University of Agriculture and Technology).

表-1 従来法 [2] と提案法における方位角の更新回数, DOA の平均絶対誤差, 4 音源の平均 Δ SDR. 更新回数と DOA 推定誤差は四捨五入前に値が低かった上位 3 条件をハイライトし, 最も低かった条件を太字で表記した.

Method	Parameters		Number of updates ($\downarrow$)				DOA error [deg] ($\downarrow$)				Δ SDR [dB] ($\uparrow$)			
	in Eq. (4)	$\psi_{\min} =$	1	0.1	0.01	0.001	1	0.1	0.01	0.001	1	0.1	0.01	0.001
Conv.	$\alpha \rightarrow 1, \beta = 0.5, \lambda = \infty$		1.94	8.00	34.56	96.56	3.52	3.25	3.04	2.96	6.86	6.83	6.83	6.83
Prop.	$\lambda = 0$		3.94	16.75	43.50	89.25	3.65	3.25	3.01	2.96	6.85	6.83	6.83	6.83
	$\alpha = 2, \beta = 0.5, \lambda = 10^0$		3.88	16.56	43.06	88.62	3.66	3.26	3.01	2.96	6.85	6.83	6.83	6.83
	$\alpha = 2, \beta = 0.5, \lambda = 10^4$		2.38	7.13	31.31	76.50	3.35	3.26	3.02	2.96	6.83	6.83	6.83	6.83
	$\alpha = 2, \beta = 0.5, \lambda = 10^8$		2.50	7.69	28.94	69.31	3.35	3.21	3.00	2.96	6.83	6.83	6.83	6.83
	$\alpha = 2, \beta = 0.5, \lambda = 10^{12}$		2.50	9.00	27.63	65.81	3.35	3.18	3.01	2.96	6.83	6.83	6.83	6.83
	$\alpha = 2, \beta = 0.5, \lambda = \infty$		2.50	9.00	27.31	68.81	3.35	3.18	3.01	2.96	6.83	6.83	6.83	6.83

にのみ値を縮小させ, それ以外は変化させない条件分岐的な更新を行っていた. そのため, 同方向へ連続更新する場合 ($d_n^{[i]} d_n^{[i-1]} > 0$) の制御や, 平均変化率の変動の大きさ ($|d_n^{[i]} d_n^{[i-1]}|$) 自体は考慮されていなかった. そこで本稿では, 方位角更新のさらなる効率化に向け, ステップサイズを平均変化率の積 $d_n^{[i]} d_n^{[i-1]}$ に応じて動的に制御する手法を提案する. 提案法におけるステップサイズの更新式を次式で与える.

$$\eta_n^{[i]} = \gamma_n^{[i]} \eta_n^{[i-1]}, \quad \gamma_n^{[i]} = \beta (\alpha/\beta)^{1/u_n^{[i]}} \quad (3)$$

ただし, γ_n はステップサイズの拡大・縮小倍率, $1/u_n \in (0, 1)$ であり, u_n は次式で与えられる.

$$u_n^{[i]} = 1 - (\log_\beta \alpha) \exp(-2\lambda d_n^{[i]} d_n^{[i-1]}) \quad (4)$$

ただし, $\alpha > 1$ と $0 < \beta < 1$ は倍率の上限と下限, $\lambda \geq 0$ は感度パラメータである. 式 (3) では, 積 $d_n^{[i]} d_n^{[i-1]}$ が正 (同方向更新) なら η_n を拡大 ($\gamma_n > 1$) して接近を加速し, 負 (方向反転) なら縮小 ($\gamma_n < 1$) して収束を促す. また, 倍率 γ_n は積の絶対値 $|d_n^{[i]} d_n^{[i-1]}|$ に応じて上限 α や下限 β へ滑らかに変化するため, 変動の激しさを反映した制御が可能となる. 特に, 積が 0 近傍となる収束間近では $\gamma_n \approx 1$ に維持され, η_n の過度な変動が抑制される.

感度パラメータ λ は, 積 $d_n^{[i]} d_n^{[i-1]}$ の変化に対する倍率 γ_n の反応の鋭さを担う. $\lambda = 0$ では積によらず一定 ($\gamma_n = 1$) となり, $\lambda = \infty$ では拡大 ($\gamma_n = \alpha$) と縮小 ($\gamma_n = \beta$) が条件分岐的に切り替わるステップ関数の挙動となる. なお, 式 (3) は $\alpha \rightarrow 1$ かつ $\lambda = \infty$ のとき, 倍率 γ_n の維持 ($\gamma_n = 1$) と縮小 ($\gamma_n = \beta$) のみを行う従来法 [2] の更新式と一致する.

4 実験

提案法の有効性を評価するため, 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, 各音源の真の DOA は $\{-50^\circ, -10^\circ, 15^\circ, 45^\circ\}$, 残響時間は 130 ms

である. 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 $\{1^\circ, 0.1^\circ, 0.01^\circ, 0.001^\circ\}$ の 4 パターンとした. 式 (1) を解くには ADMM アルゴリズム [3] を用いた. さらに, 式 (3), 式 (4) では, ステップサイズの初期値 $\eta_n^{[0]}$ は 200, 拡大・縮小倍率の上限 α は 2, 下限 β は 0.5 とした. 感度パラメータ λ については $\{0, 10^0, 10^4, 10^8, 10^{12}, \infty\}$ の 6 パターンで比較を行い, 併せて従来法 [2] ($\alpha \rightarrow 1, \beta = 0.5, \lambda = \infty$) との比較も行った.

実験結果を表-1 に示す. まず, 収束条件が緩い $\psi_{\min} = 1^\circ$ では, 提案法は方位角の更新回数が増加したものの, ステップサイズの拡大による迅速な接近によって DOA 推定誤差が減少した. 次に, より厳しい収束条件 ($\psi_{\min} \leq 0.01^\circ$) では, 提案法は高い DOA 推定精度を維持しつつ, λ の設定に応じて更新回数を削減できた. これは, 平均変化率の積に応じた動的なステップサイズ制御により, 迅速な追従と安定した収束が両立されたためである. 特に, 従来の条件分岐的な切り替えに相当する極限 ($\lambda = \infty$) よりも, 有限の λ (本実験では $\lambda = 10^{12}$) で更新回数が最小となり, 滑らかな動的制御の優位性が示された. なお, Δ SDR は条件によらず同等であり, 従来と同等の分離性能が維持されていることを確認した.

5 むすび

本稿では, 方位角更新におけるステップサイズを制御する手法を提案した. 実験の結果, ステップサイズの動的な拡大・縮小が, DOA 推定に必要な方位角の更新回数の削減に有効であることを示した.

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, "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," 日本音響学会講演論文集, pp. 149–150 (2026.3).
- [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).