

# 幾何学的制約付き独立ベクトル分析を用いたオンライン指向性音声強調の Iterative Source Steering による高速化\*

☆後藤加奈 (筑波大), 上田哲也 (早稲田大), 李莉 (NTT),  
山田武志 (筑波大), 牧野昭二 (早稲田大, 筑波大)

## 1 はじめに

マイクロホン数が音源数と等しい場合に目的音声  
を抽出する有用な手法として, 独立成分分析 (Independent Component Analysis: ICA) を含むブライ  
ンド音源分離 (Blind Source Separation: BSS) が広  
く研究されている. 独立ベクトル分析 (Independent  
Vector Analysis: IVA) [1, 2] は周波数成分全体を球状  
の多変量分布に従う多変量変数としてモデル化する  
ことで音源分離問題と周波数間順序問題を同時に解  
決する手法の一つである. 近年では補助関数法を用い  
た IVA (Auxiliary function-based IVA: AuxIVA) [3]  
が提案されている. AuxIVA は Iterative Projection  
(IP) [4] を用いて分離フィルタ更新を行い, ステップ  
サイズパラメータが不要でコスト関数が高速収束する  
といった実用化のために好ましい利点を持つ. 一方,  
電話会議システムや補聴器などのリアルタイムアプ  
リケーションにとってオンライン処理は重要であり,  
時間的に逐次に観測される信号のみで分離するオン  
ラインアルゴリズムの開発は不可欠である. AuxIVA  
をフレームごとに更新する Online AuxIVA [5] はそ  
の一つの成功例である.

これらの BSS 手法は音源間の独立性に基づいて分離  
を行うため話者情報や事前学習が不要であるという利  
点がある. 一方で出力チャンネルの順序に任意性がある  
ため, 話者情報や空間情報を用いて分離後に目的音声  
を選択するための後処理が必要となる. このような出力  
順序問題と音源分離問題を同時最適化するためのア  
プローチとして, 音源到来方向 (Direction Of Arrival:  
DOA) やセンサ位置といった空間情報を与えて分離  
行列を求める幾何学的制約付き BSS (Geometrically  
Constrained BSS: GC-BSS) が研究されている [6-8].  
中でもベクトル座標降下法に基づく幾何学的制約付  
き AuxIVA (GC-AuxIVA with Vectorwise Coordinate  
Descent: 本稿では GC-AuxIVA-VCD として説明す  
る) [6] は VCD [7] のアイデアを取り入れて導出した  
更新アルゴリズムによって AuxIVA の利点を保持し  
つつ, 出力順序問題を効率的に解決する手法である.  
さらにそのオンライン手法も提案されており, 有効性  
が確認されている [9].

しかし AuxIVA や GC-AuxIVA-VCD における更  
新則は, 逆行列計算が含まれるため大きな計算コス  
トを要し数値計算に不安定さを招くなどの問題があ  
る. このような問題を克服するため, Iterative Source  
Steering (ISS) と呼ばれる逆行列演算不要で計算コ  
ストが小さい新たな更新アルゴリズムを用いた Aux-  
IVA (AuxIVA-ISS) が近年提案され [10], IP に基づく  
AuxIVA と比較し実行時間を削減したことが示されて  
いる. また ISS の更新規則は混合行列の列ベクトルを

更新することと同義であるため, 移動音源と固定音源  
が混合された状況では, 分離行列のうち移動音源に対  
応するベクトルのみを逐次更新することでさらに効率  
的に計算出来ることが示されている [11]. GC-BSS 分  
野では, 我々は計算量の削減のため ISS に基づく GC-  
AuxIVA のオフラインアルゴリズム (GC-AuxIVA-  
ISS) を提案し, 短い演算時間で従来法の GC-AuxIVA-  
VCD と同等かそれ以上の性能を達成した [12]. 更に  
実用化に近付けるため, 本稿では, GC-AuxIVA-ISS  
のオンラインアルゴリズムを提案する. また現実で  
は移動音源の DOA が既知であると仮定することは  
強い制限となるため, Multiple Signal Classification  
(MUSIC) [13] による DOA 推定を取り入れた実験を  
行い, 指向性音声強調における提案手法の性能を評  
価する.

## 2 従来手法

### 2.1 Offline GC-AuxIVA-VCD

$I$  個のマイクロホンで  $J$  個の音源から到来する  
信号を観測する場合を考える. 各時間周波数ビン  
における観測信号と推定音源の短時間フーリエ変  
換 (short-time Fourier transform: STFT) 係数を  
それぞれ  $\mathbf{x}_{fn} = [x_{1fn}, \dots, x_{Ifn}]^T \in \mathbb{C}^I$ ,  $\mathbf{y}_{fn} =$   
 $[y_{1fn}, \dots, y_{Jfn}]^T \in \mathbb{C}^J$  とする. ただし  $f = 1, \dots, F$   
と  $n = 1, \dots, N$  はそれぞれ周波数ビンと時間フレ  
ームのインデックスであり,  $(\cdot)^T$  は行列の転置である.  
本稿では  $I = J$  の場合を扱う. STFT 窓の長さが音  
源とマイクロホン間のインパルス応答よりも十分に  
長い場合, 観測信号と推定音源の関係は時不変瞬時  
混合モデルを用いて次のように表すことができる:

$$\mathbf{y}_{fn} = \mathbf{W}_f \mathbf{x}_{fn}, \quad (1)$$

ここで  $\mathbf{W}_f = [\mathbf{w}_{1f}, \dots, \mathbf{w}_{Jf}]^H$  は,  $\mathbf{w}_{jf} =$   
 $[w_{1jf}, \dots, w_{Ijf}]^T$  で表される分離フィルタを要素と  
する  $I \times I$  の分離行列,  $(\cdot)^H$  はエルミート転置を表す.

IVA における分離行列  $\mathcal{W} = \{\mathbf{W}_f\}_f$  は次の目的関  
数を最小化することにより推定される:

$$\mathcal{L}_{\text{IVA}}(\mathcal{W}) = \sum_{j=1}^J \mathbb{E}[G(\mathbf{y}_{jn})] - \sum_{f=1}^F \log |\det \mathbf{W}_f|, \quad (2)$$

ここで  $\mathbb{E}[\cdot]$  は期待値演算子を表す.  $\mathbf{y}_{jn} =$   
 $[y_{j1n}, \dots, y_{jFn}]^T \in \mathbb{C}^F$  は音源単位でのベクト  
ル表現である.  $G(\mathbf{y}_{jn})$  は  $G(\mathbf{y}_{jn}) = -\log p(\mathbf{y}_{jn})$  の  
関係を持つコントラスト関数であり,  $p(\mathbf{y}_{jn})$  は  $j$  番  
目の音源の  $n$  番目のフレームにおける多変量確率密  
度関数である. 一般的に, コントラスト関数は以下

\*Fast online algorithm for directional speech enhancement using geometrically constrained independent vector analysis with iterative source steering, by GOTO, Kana (University of Tsukuba), UEDA, Tetsuya (Waseda University), LI, Li (NTT), YAMADA, Takeshi (University of Tsukuba), MAKINO, Shoji (Waseda University, University of Tsukuba)

の式のような球状多変量分布が用いられる。

$$G(\mathbf{y}_{jn}) = G_R(r_{jn}), \quad (3)$$

$$r_{jn} = \|\mathbf{y}_{jn}\|_2 = \sqrt{\sum_f |y_{jfn}|^2} = \sqrt{\sum_f |\mathbf{w}_{jf}^H \mathbf{x}_{fn}|^2}. \quad (4)$$

補助関数法に基づき、元の目的関数の代わりに最適化する補助関数を以下のように設計する：

$$\begin{aligned} \mathcal{L}_{\text{IVA}}(\mathcal{W}) &\leq \mathcal{L}_{\text{AuxIVA}}(\Sigma, \mathcal{W}) \\ &= \frac{1}{2} \sum_{f=1}^F \sum_{j=1}^J \mathbf{w}_{jf}^H \Sigma_{jf} \mathbf{w}_{jf} - \sum_{f=1}^F \log |\det \mathbf{W}_f|. \end{aligned} \quad (5)$$

ここで、 $\Sigma = \{\Sigma_{jf}\}_{jf}$  であり、 $\Sigma_{jf}$  は次のように表される加重共分散行列である：

$$\Sigma_{jf} = \sum_n \varphi(r_{jn}) \mathbf{x}_{fn} \mathbf{x}_{fn}^H. \quad (6)$$

ここで  $\varphi(r_{jn}) = G'_R(r_{jn})/r_{jn}$  であり、 $(\cdot)'$  は微分演算子を表す。

また、IVA で推定された  $j$  番目の分離フィルタの遠距離場応答を  $\theta$  方向に制限する幾何学的制約を考えると、次のような目的関数が得られる：

$$\mathcal{L}_{\text{GC}}(\mathcal{W}) = \sum_{f=1}^F \sum_{j=1}^J \sum_{\theta \in \Theta} \lambda_{j\theta} |\mathbf{w}_{jf}^H \mathbf{d}_{f\theta} - c_{j\theta}|^2. \quad (7)$$

ここで、 $\mathbf{d}_{f\theta}$  は  $\theta$  方向を指すステアリングベクトル、 $c_{j\theta}$  は非負の値を持つ制約、 $\lambda_{j\theta} \geq 0$  は制約を重み付けするパラメータである。  $c_{j\theta} = 1$  の場合、空間フィルタは  $\theta$  方向の目的信号を維持するように導かれる。これは従来の遅延和法 (Delay and Sum: DS) ビームフォーマと等価である。  $c_{j\theta}$  の値が小さい場合は、 $\theta$  方向の音源を抑制し他の全ての音源を保存することを目的とした、 $\theta$  方向への空間的なヌルを形成する。幾何学的制約項は容易に最適化可能な線形形式であるため、補助関数を設計せずに最適化を行うことが可能である。

従って、GC-AuxIVA-VCD で最適化したい補助関数は以下のように表される：

$$\mathcal{L}(\Sigma, \mathcal{W}) = \mathcal{L}_{\text{AuxIVA}}(\Sigma, \mathcal{W}) + \mathcal{L}_{\text{GC}}(\mathcal{W}). \quad (8)$$

ベクトル座標降下法 (Vectorwise Coordinate Descent: VCD) [7] を利用し、導出される GC-AuxIVA-VCD の更新規則は以下のようにまとめられる：

$$\mathbf{u}_{jf} = \mathbf{D}_{jf}^{-1} \mathbf{W}_f^{-1} \mathbf{e}_j, \quad (9)$$

$$\hat{\mathbf{u}}_{jf} = \mathbf{D}_{jf}^{-1} \sum_{\theta \in \Theta} \lambda_{j\theta} c_{j\theta} \mathbf{d}_{f\theta}, \quad (10)$$

$$\mathbf{h}_{jf} = \mathbf{u}_{jf}^H \mathbf{D}_{jf} \mathbf{u}_{jf}, \quad (11)$$

$$\hat{\mathbf{h}}_{jf} = \mathbf{u}_{jf}^H \mathbf{D}_{jf} \hat{\mathbf{u}}_{jf}, \quad (12)$$

$$\mathbf{w}_{jf} = \begin{cases} \frac{1}{\sqrt{\hat{\mathbf{h}}_{jf}}} \mathbf{u}_{jf} + \hat{\mathbf{u}}_{jf} \quad (\hat{\mathbf{h}}_{jf} = 0), \\ \frac{\hat{\mathbf{h}}_{jf}}{2\hat{\mathbf{h}}_{jf}} \left[ -1 + \sqrt{1 + \frac{4\hat{\mathbf{h}}_{jf}}{|\hat{\mathbf{h}}_{jf}|^2}} \right] \mathbf{u}_{jf} + \hat{\mathbf{u}}_{jf} \quad (\text{o.w.}). \end{cases} \quad (13)$$

ここで  $\mathbf{D}_{jf} = \Sigma_{jf} + \sum_{\theta \in \Theta} \lambda_{j\theta} \mathbf{d}_{f\theta} \mathbf{d}_{f\theta}^H$ 、 $\mathbf{e}_j$  は  $J \times J$  単位行列の  $j$  番目のベクトルである。導出の詳細は [6,7] に記載されている。また、本更新規則は全ての  $j$  と  $\theta$  について  $\lambda_{j\theta} = 0$  の時、IP に基づく AuxIVA の更新規則と同等になる。

## 2.2 Online GC-AuxIVA-VCD

Offline GC-AuxIVA-VCD では全ての時間フレームでの観測信号を用いて共分散行列  $\Sigma_{jf}$  が計算される。しかしオンライン処理の場合は現在までの観測信号のみしか得られないため、以下のように共分散行列を各時間フレーム  $n$  で回帰的に近似計算する：

$$\Sigma_{jfn} = \alpha \Sigma_{jfn-1} + (1 - \alpha) \varphi(r_{jn}) \mathbf{x}_{fn} \mathbf{x}_{fn}^H. \quad (14)$$

ここで  $0 \leq \alpha < 1$  は忘却係数である。ただし、 $r_{jn}$  は時変な分離フィルタ  $\mathbf{w}_{jfn}$  を用いて計算され、該当フレーム  $n$  の分離フィルタ  $\mathbf{w}_{jfn}$  はフレーム  $n-1$  において更新されたものが用いられる。Online GC-AuxIVA-VCD の更新規則は、Offline GC-AuxIVA-VCD での  $\Sigma_{jf}$  の更新を式 (6) から式 (14) に置き換えることで得られる [9]。

## 3 提案手法: Online GC-AuxIVA-ISS

前節で紹介した Online GC-AuxIVA-VCD は逆行列演算によって計算コストが大きくなるという問題がある。一方で BSS 分野では ISS [10] に基づく Online AuxIVA-ISS [11] が提案されている。ISS は逆行列演算を含まないため計算コストが小さく計算が安定する利点がある。また GC-BSS 分野においても、ISS の利点を持つ GC-AuxIVA-ISS と呼ばれるオフライン更新アルゴリズムが提案されている [12]。そこで本節では Online GC-AuxIVA-VCD の利点と ISS の計算上の利点を併せ持つ手法を提案するため、GC-AuxIVA-ISS をオンライン処理に拡張する。

ISS では分離フィルタの代わりに  $\mathbf{v}_{jf} = [v_{1jf}, \dots, v_{Ijf}]^T \in \mathbb{C}^J$  で定義される補助変数を更新することで分離行列を以下のように最適化する：

$$\mathbf{W}_{fn} \leftarrow \mathbf{W}_{fn} - \mathbf{v}_{jf} \mathbf{w}_{jfn}^H. \quad (15)$$

Online GC-AuxIVA-VCD の目的関数に式 (15) を代入し  $\mathbf{v}_{jf}$  について解くと、以下の更新式が得られる：

$$v_{ijf} = \frac{\mathbf{w}_{ifn} \Sigma_{ifn} \mathbf{w}_{jfn}^H + 2 \sum_{\theta \in \Theta} \lambda_{i\theta} g_{j\theta}^* (g_{i\theta} - c_{i\theta})}{\mathbf{w}_{jfn} \Sigma_{ifn} \mathbf{w}_{jfn}^H + 2 \sum_{\theta \in \Theta} \lambda_{i\theta} |g_{j\theta}|^2} \quad (\forall i \neq j), \quad (16)$$

$$v_{jff} = \begin{cases} 1 - p_{jf}^{-1/2} \quad (q_{jff} = 0), \\ 1 - q_{jf}^* \frac{|q_{jff}| + \sqrt{|q_{jff}|^2 + p_{jf}}}{p_{jf} |q_{jff}|} \quad (\text{o.w.}), \end{cases} \quad (17)$$

ここで  $g_{jff} = \mathbf{w}_{jfn}^H \mathbf{d}_{f\theta}$ 、 $p_{jf} = \mathbf{w}_{jfn} \Sigma_{jfn} \mathbf{w}_{jfn}^H + 2 \sum_{\theta \in \Theta} \lambda_{j\theta} |g_{j\theta}|^2$ 、 $q_{jff} = \sum_{\theta \in \Theta} \lambda_{j\theta} c_{j\theta} g_{j\theta}$  である。また前節と同様に、この更新規則は全ての  $j$  と  $\theta$  について  $\lambda_{j\theta} = 0$  の時、Online AuxIVA-ISS の更新規則と同等になる。

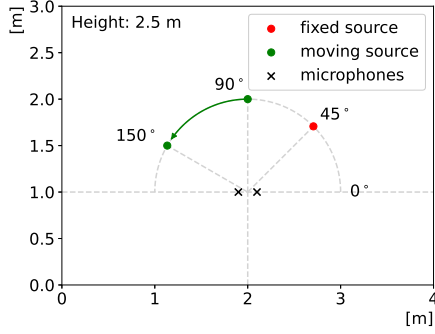


Fig. 1: Layout of sound sources and microphones.

## 4 評価実験

### 4.1 条件

Online GC-AuxIVA-ISS の指向性音声強調における有効性を評価するため、シミュレーション実験を行った。本実験では、ATR デジタル音声データベースのセット B に収録されている、全 503 文の音素バランス文の男性 3 話者、女性 3 話者計 6 話者分のデータを使用した。このデータセットの中からランダムに選択した 2 人の異なる話者の発話に対し、鏡像法で合成した室内のインパルス応答を畳み込むことで、2 チャンネル、信号長 60 秒の観測信号を 20 パターン作成した。混合音の合成には `signal generator`<sup>1</sup> を使用した。残響時間 ( $RT_{60}$ ) は 200 ms, サンプリング周波数は 16 kHz とした。音源とマイクロホンの配置を図 1 に示す。目的音は固定音源であり、干渉音は最初の 20 秒間は  $90^\circ$  に固定、次の 20 秒間は  $90^\circ$  から  $150^\circ$  方向へと円弧上を移動し、最後の 20 秒間は  $150^\circ$  に固定である。STFT はハン窓を使用し、窓長と移動長はそれぞれ 64, 32 ms とした。加重共分散行列  $\Sigma_{j f_0}$  と分離行列の  $\mathbf{W}_{f_0}$  はいずれも単位行列で初期化した。忘却係数  $\alpha$  は 0.99 とした。

本実験では Online AuxIVA-ISS (oAuxIVA-ISS) [11], Online GC-AuxIVA-VCD (oGC-AuxIVA-VCD) [9], 提案手法 Online GC-AuxIVA-ISS (oGC-AuxIVA-ISS) の 3 手法を比較する。oGC-AuxIVA-VCD と oGC-AuxIVA-ISS については干渉音方向にヌルを生成する幾何学的制約を採用し、 $\theta$  が干渉音方向の場合全ての  $j$  において  $\lambda_{j\theta} = \lambda_{\text{null}} > 0$ , それ以外の  $\theta$  については  $\lambda_{j\theta} = 0$ , また全ての  $j$  と  $\theta$  において  $c_{j\theta} = 0$  とした。評価指標は分離信号の品質 (Signal-to-Distortion Ratio: SDR), 実行時間を用いる。SDR は値が高い程高い分離性能を示す。また、oGC-AuxIVA-VCD と oGC-AuxIVA-ISS については出力順序誤りの発生率も評価する。

### 4.2 MUSIC 法を用いた実験システム

本実験では事前学習や実測等を行わずに幾何学的制約で用いられるステアリングベクトルを得るため、音源の平面波到来仮定に基づいた算出を行う。しかし移動音源の DOA を正確に知ることは出来ないため、本実験では MUSIC 法 [13] で逐次的に移動音源の DOA を推定する。

$I > J$  である場合、MUSIC 法によって推定される  $\theta$  方向に対する空間スペクトル  $P_{f\theta}$  は以下の式で与え

られる:

$$P_{f\theta} = \frac{1}{\sum_{i=J+1}^I |\mathbf{d}_{f\theta}^H \mathbf{u}_{if}|^2}. \quad (18)$$

ここで  $\mathbf{u}_{if}$  ( $i = J+1, \dots, I$ ) は以下を満たす固有ベクトルである:

$$\mathbf{R}_f \mathbf{u}_{if} = \mu_{if} \mathbf{u}_{if}. \quad (19)$$

ここで  $\mathbf{R}_f$  は観測信号の空間共分散行列であり、 $\mu_{1f}, \dots, \mu_{if}, \dots, \mu_{If}$  は降順に並べられた  $\mathbf{R}_f$  の固有値とする。本実験では各分離手法によって推定された分離信号を projection back 法 [14] でスケール補正することで、1 つの音源に対しマイクロホン数の観測信号を疑似的に生成し、入力として用いた:

$$\tilde{\mathbf{y}}_{jfn} = \mathbf{W}_f^{-1} \mathbf{e}_j y_{jfn}, \quad (20)$$

また空間共分散行列は  $\mathbf{R}_f = \mathbb{E}[\tilde{\mathbf{y}}_{jfn} \tilde{\mathbf{y}}_{jfn}^H]$  となる。信号部分空間と雑音部分空間の直交性より  $\mathbf{d}_{f\theta}$  が実際のステアリングベクトルと一致する場合に  $P_{f\theta}$  はピークとなる。本実験では、各周波数帯域でピークとなる  $\theta$  を算出し、それらを平均した角度を推定 DOA とした。

oGC-AuxIVA-VCD と oGC-AuxIVA-ISS について未知である干渉音の DOA を得るため、本実験では 1 フレーム分の観測信号を用いて MUSIC による DOA 推定を行う “MUSIC normal”, 1 フレーム分の観測信号を用いて MUSIC による DOA 推定を行ったあとに最新 5 フレームの結果を平均する “MUSIC smooth”, 5 フレーム毎に 5 フレーム分の観測信号を用いて MUSIC による DOA 推定を行う “MUSIC block”, また性能の上限を示す例として正しい DOA を与える “Correct” の 4 つのアプローチを比較する。5 フレーム間の移動音源の移動距離はおよそ  $0.4^\circ$  ( $0.7 \text{ cm}$ ) であり、十分に固定音源としての仮定が成り立つ。式 (18) の計算は `pyroomacoustics`<sup>2</sup> を使用した。

### 4.3 結果

図 2 は各手法によって強調された目的音源 (固定音源) の 2 秒毎の平均 SDR を示している。幾何学的制約の重み  $\lambda_{\text{null}}$  は、出力順序誤りが最も少ない中で最も高い SDR を示した値を採用した。oGC-AuxIVA-VCD と oGC-AuxIVA-ISS の両手法は  $\lambda_{\text{null}}$  の値が適切であれば出力順序誤りが起こらないことを確認した。干渉音の移動直後に大きく SDR が減少しているが、幾何学的制約を用いた oGC-AuxIVA-VCD と oGC-AuxIVA-ISS はブラインドな oAuxIVA-ISS より早く分離性能が改善している。一方で図 3 より、“MUSIC normal” や “MUSIC block” は “MUSIC smooth” に比べ各フレームにおける推定 DOA に大きくばらつきが生じ、これらにおいて提案手法 oGC-AuxIVA-ISS はベースライン手法 oGC-AuxIVA-VCD の SDR スコアを僅かに下回った。これより DOA 推定に大きく誤差が生じる場合は特に提案手法の性能低下の原因となると考えられ、その解決は今後の課題である。

表 1 は 60 秒の観測信号を分離した実行時間を示している。いずれの DOA 推定手法においても提案手法

<sup>1</sup><https://www.audiolabs-erlangen.de/fau/professor/habets/software/signalgenerator>

<sup>2</sup><https://pyroomacoustics.readthedocs.io/en/pyipi-release/>

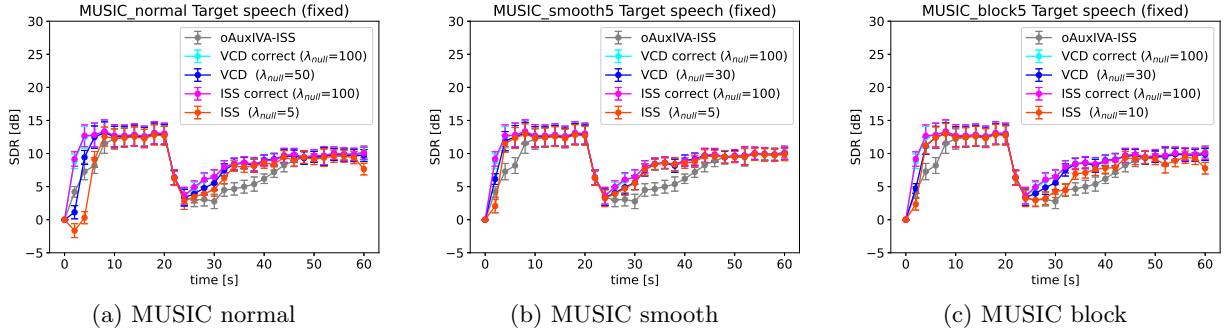


Fig. 2: Average SDR and SIR of the target source (fixed) enhanced by each method in every 2 s. Error bar denotes the  $1.96 \times$  standard error in each time. VCD and ISS in the legend denote oGC-AuxIVA-VCD and oGC-AuxIVA-ISS (proposed), respectively, and  $\lambda_{nuII}$  denotes the weight of the geometric constraint. For oGC-AuxIVA-VCD and oGC-AuxIVA-ISS, the accuracy of output signal order were 100%.

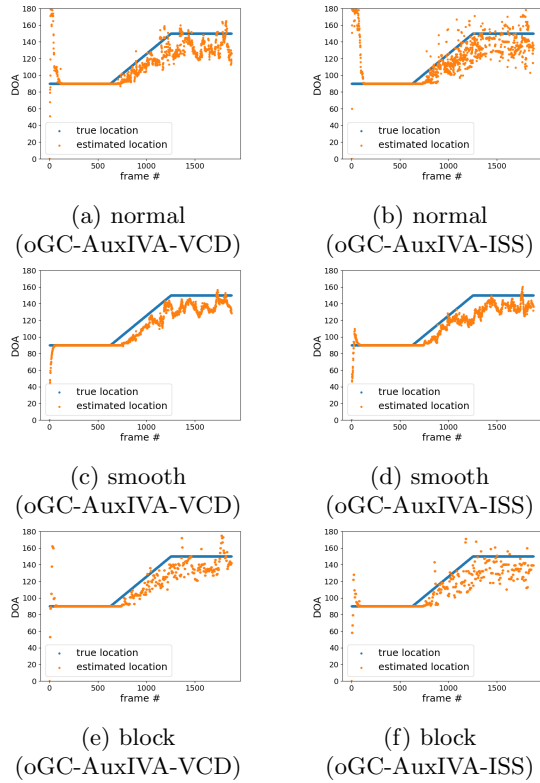


Fig. 3: DOA of the moving source estimated by MUSIC. Blue lines represent the correct DOA.

はベースライン手法より短い実行時間で分離できていることが示された。

## 5 おわりに

本稿では計算コストの少ないオンライン GC-BSS 手法を提案した。提案手法は ISS によって、従来手法の欠点であった逆行列演算により計算コストが大きくなる問題と計算が不安定になる問題を解消できる。指向性音声強調実験により、提案手法は従来手法とほぼ同等の強調性能を維持しながら実行時間の大幅な削減を達成した。また、MUSIC 法による DOA 推定を用いた GC-BSS の実験システムは、BSS による音源分離よりも移動音源の抑圧に高い性能を示すことが示された。

Table 1: Average runtime for 60 second signals [s].

DOA estimation	oGC-AuxIVA-VCD [6]	oGC-AuxIVA-ISS (proposed)
Correct	25.25	6.24
MUSIC normal	75.32	56.04
MUSIC smooth	73.96	55.99
MUSIC block	35.58	16.16

謝辞 本研究は科研費 19H04131 の助成を受けた。

## 参考文献

- [1] T. Kim *et al.*, in *Proc. ICA*, 2006, pp. 165–172.
- [2] A. Hiroe, in *Proc. ICA*, 2006, pp. 601–608.
- [3] N. Ono, in *Proc. WASPAA*, 2011, pp. 189–192.
- [4] N. Ono *et al.*, in *Proc. LVA/ICA*, 2010, pp. 165–172.
- [5] T. Taniguchi *et al.*, in *Proc. HSCMA*, 2014, pp. 107–111.
- [6] L. Li *et al.*, in *Proc. ICASSP*, 2020, pp. 846–850.
- [7] Y. Mitsui *et al.*, in *Proc. ICASSP*, 2018, pp. 746–750.
- [8] K. Goto *et al.*, in *Proc. APSIPA*, 2020, pp. 858–862.
- [9] L. Li *et al.*, in *Proc. Interspeech*, 2020, pp. 61–65.
- [10] R. Scheibler *et al.*, in *Proc. ICASSP*, 2020, pp. 236–240.
- [11] T. Nakashima *et al.*, in *Proc. Spring Meeting of ASJ*, 2022, pp. 185–188.
- [12] K. Goto *et al.*, in *Proc. EUSIPCO*, 2022.
- [13] R. Schmidt, *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.
- [14] N. Murata *et al.*, *Neurocomputing*, vol. 41, no. 1–4, pp. 1–24, 2001.