

近接分離最適化によるブラインド音源分離*

○矢田部浩平 (早大), 北村大地 (東大)

1 まえがき

N 個の音源が M 本のマイクロホンによって観測される状況を、時間周波数領域において

$$\mathbf{x}[t, f] \approx A[f]\mathbf{s}[t, f] \quad (1)$$

と近似する。ただし、 $\mathbf{x} = [x_1, x_2, \dots, x_M]^T$ は観測信号、 $\mathbf{s} = [s_1, s_2, \dots, s_N]^T$ は復元されるべき音源信号、 $A[f]$ は $M \times N$ の混合行列、 t と f はそれぞれ時間と周波数のインデックスを表す。ブラインド音源分離問題では、観測信号 \mathbf{x} のみから、(定数倍を除いて)元の音源信号 \mathbf{s} を復元することを考える。この問題は、優決定な場合 ($M \geq N$)、 $N \times M$ の分離行列 $W[f]$ ($A[f]$ の左逆行列) を求める問題として定式化され、

$$W[f]\mathbf{x}[t, f] \approx W[f]A[f]\mathbf{s}[t, f] = \mathbf{s}[t, f] \quad (2)$$

のように分離信号を得ることが多い。本稿でもこの定式化に則り、ブラインド音源分離を $W[f]$ の推定問題として考えるが、以降は簡単のため、 $A[f]$ および $W[f]$ を正方行列 ($N = M$) の場合に限って議論する (ただし本稿の提案は $M \neq N$ の場合も含められる)。

分離行列 $W[f]$ の推定には、音源信号の統計的独立性の仮定に基づく手法が広く用いられている。例えば、独立成分分析 (ICA) を周波数領域で適用した FDICA や、そのパーミュテーション問題に対処する手法である独立ベクトル分析 (IVA)、さらに近年では音源モデルに非負値行列因子分解 (NMF) を導入した独立低ランク行列分析 (ILRMA) [1, 2] などが盛んに研究されている。これらの手法は、音源信号の事前知識 (非ガウス性、共起性、低ランク良近似性など) に基づいて分離を実現しており、音源信号のモデルの良し悪しが手法の良し悪しを特徴付けていると解釈できる (式 (3) および式 (4)~(6) 参照)。すなわち、より良い音源モデルを導入することができれば、より良い音源分離手法を実現できる可能性がある。

より良い音源モデルを探求するためには、異なる音源モデル同士の比較が重要であり、従って幅広い音源モデルに対応可能なアルゴリズムが求められる。しかし、近年主流のアルゴリズムである補助関数法 [2-4] は、音源モデルに合わせて特別に補助関数を設計する必要があり、新たな音源モデルに対応するためには、経験則に基づく補助関数の発見が要求される。そのように問題ごとに特別に設計されたアルゴリズムは、

計算上有利な場合が多く、高速化や実用化を考えるとときには非常に有効であるが、一方で、様々な音源モデルを試しながら探求する作業に対しては、アルゴリズムの発見に時間を要する場合があります、必ずしも適しているとは言えない。

そこで本研究では、近接分離最適化手法 [5-8] に基づき、幅広い音源モデルを統一的に扱うことが可能なアルゴリズムを提案する [9]。提案手法は、近接作用素が計算可能な音源モデルであれば、どのようなモデルでも統一的に扱えるアルゴリズムとなっている。反復式中の近接作用素を変更するだけで各モデルに対応可能であり、プログラミングの手間をかけずに各モデルを試すことができるので、より良い音源モデルの発見に寄与することが期待される。

2 独立性に基づく音源分離

上述の通り、優決定の音源分離問題は $M \times M$ の分離行列たち $\{W[f]\}_{f=1}^F$ を求める問題として定式化され、その多くは以下の形式の最小化問題に帰着する。

$$\text{Minimize}_{\{W[f]\}_{f=1}^F} \mathcal{P}(W[f]\mathbf{x}[t, f]) - \sum_{f=1}^F \log |\det(W[f])| \quad (3)$$

ただし、 \mathcal{P} は音源モデルから導かれる実数値関数であり、例えば代表的な手法は、適当な定数 C を用いて以下のように書くことができる。

Laplace 分布に基づく FDICA

$$\mathcal{P}(\mathbf{y}[t, f]) = C \|\mathbf{y}[t, f]\|_1 = C \sum_{m=1}^M \sum_{t=1}^T \sum_{f=1}^F |y_m[t, f]| \quad (4)$$

球対称 Laplace 分布に基づく IVA

$$\mathcal{P}(\mathbf{y}[t, f]) = C \|\mathbf{y}[t, f]\|_{2,1} = C \sum_{m=1}^M \sum_{t=1}^T \left(\sum_{f=1}^F |y_m[t, f]|^2 \right)^{\frac{1}{2}} \quad (5)$$

板倉斎藤 NMF に基づく ILRMA

$$\mathcal{P}(\mathbf{y}[t, f]) = C \sum_{m=1}^M \mathcal{D}_R(y_m[t, f]) \quad (6)$$

ただし、 \mathcal{D}_R は板倉斎藤 NMF に基づくパワースペクトログラムの低ランク性の指標である。

$$\mathcal{D}_R(y_m[t, f]) = \min_{\varphi_{f,r}^{[m]} \geq 0, \psi_{r,t}^{[m]} \geq 0} \sum_{t=1}^T \sum_{f=1}^F \left(\frac{|y_m[t, f]|^2}{\sum_{r=1}^R \varphi_{f,r}^{[m]} \psi_{r,t}^{[m]}} + \log \sum_{r=1}^R \varphi_{f,r}^{[m]} \psi_{r,t}^{[m]} \right) \quad (7)$$

*Blind source separation via proximal splitting algorithm. By Kohei YATABE (Waseda University) and Daichi KITAMURA (The University of Tokyo).

式 (3) の視点で各手法を眺めると、実数値関数 \mathcal{P} の違いが、各手法の分離結果の違いをもたらしていることが直ちにわかる。そこで本研究では、 \mathcal{P} の形式に関わらず、その近接作用素さえ計算できれば、式 (3) の形の問題を統一的に扱えるアルゴリズムを提案する。

3 提案手法

本稿の提案は、主双対分離アルゴリズムに基づいており、従って、まずは主双対分離手法の説明をした後、式 (3) の問題に適用する方法について述べる。

3.1 主双対分離アルゴリズム

g と h を下半連続な真凸関数として、行列 L の絡んだ最適化問題を考える。

$$\underset{\mathbf{w}}{\text{Minimize}} \quad g(\mathbf{w}) + h(L\mathbf{w}) \quad (8)$$

この問題は、以下の反復によって定義される主双対分離アルゴリズム [7] によって解くことができる。

$$\begin{cases} \tilde{\mathbf{w}} = \text{prox}_{\mu_1 g}[\mathbf{w}^{[k]} - \mu_1 \mu_2 L^H \mathbf{y}^{[k]}] \\ \mathbf{z} = \mathbf{y}^{[k]} + L(2\tilde{\mathbf{w}} - \mathbf{w}^{[k]}) \\ \tilde{\mathbf{y}} = \mathbf{z} - \text{prox}_{h/\mu_2}[\mathbf{z}] \\ (\mathbf{w}^{[k+1]}, \mathbf{y}^{[k+1]}) = \alpha(\tilde{\mathbf{w}}, \tilde{\mathbf{y}}) + (1 - \alpha)(\mathbf{w}^{[k]}, \mathbf{y}^{[k]}) \end{cases} \quad (9)$$

ただし、 $\mu_1 > 0$ と $\mu_2 > 0$ は後に述べるステップサイズパラメータ、 α は $(0, 2)$ から自由に選べる加減速パラメータであり、 $\alpha = 1$ は加速なし、 $\alpha > 1$ は加速、 $\alpha < 1$ は減速を意味する。この近接分離手法の重要な特徴の 1 つは、各関数が近接作用素 [6]

$$\text{prox}_{\mu g}[\mathbf{y}] = \arg \min_{\mathbf{z}} \left[g(\mathbf{z}) + \frac{1}{2\mu} \|\mathbf{y} - \mathbf{z}\|_2^2 \right] \quad (10)$$

を通じて最小化されることである。近接作用素 prox は、式 (4) および (5) にあるような微分不可能な点を含む関数や、式 (3) に含まれる $-\log$ のように非有限関数に対応することができる。すなわち、微分不可能性や非有限性など関数の性質に起因する困難さが $(\|\cdot\|_2^2)$ との極小畳み込みにより近接作用素を通じて解決されており、従って式 (10) の計算量のみが解法の困難さとなる。幸いなことに、上で述べた音源分離手法に現れる近接作用素は、以下のように解析的に解くことができ、非常に効率的に計算可能である。

$$\text{prox}_{-\mu \log}[y] = \frac{y + \sqrt{y^2 + 4\mu}}{2} \quad (11)$$

$$\left(\text{prox}_{\mu \|\cdot\|_1}[\mathbf{y}] \right)_m[t, f] = \left(1 - \frac{\mu}{|y_m[t, f]|} \right)_+ y_m[t, f] \quad (12)$$

$$\left(\text{prox}_{\mu \|\cdot\|_{2,1}}[\mathbf{y}] \right)_m[t, f] = \left(1 - \frac{\mu}{\left(\sum_{f=1}^F |y_m[t, f]|^2 \right)^{\frac{1}{2}}} \right)_+ y_m[t, f] \quad (13)$$

ただし、 $(\cdot)_+ = \max\{0, \cdot\}$ である。

3.2 音源分離問題の再定式化

式 (3) は主双対分離アルゴリズムを直接適用できる形式になっていないので、問題を再定式化する必要がある。まず、近接作用素を考えるために、 $|\det(W[f])| = \prod_{m=1}^M \sigma_m(W[f])$ なる関係式を利用して

$$\underset{\{W[f]\}_{f=1}^F}{\text{Minimize}} \quad \mathcal{P}(W[f]\mathbf{x}[t, f]) - \sum_{f=1}^F \sum_{m=1}^M \log \sigma_m(W[f]) \quad (14)$$

と書き直す¹。ただし、 $\sigma_m(W[f])$ は $W[f]$ の m 番目の特異値を表す。

次に、最適化変数 $\{W[f]\}_{f=1}^F$ をベクトル化する。 \mathcal{V} を行列をベクトルに対応付けるベクトル化作用素

$$\mathcal{V}(W[f]) = [W_{1,1}[f], \dots, W_{1,M}[f], W_{2,1}[f], \dots, W_{M,M}[f]]^T \quad (15)$$

とし、 \mathcal{M} をその逆の行列化作用素とする。 $\{W[f]\}_{f=1}^F$ を全てまとめた $M^2 F$ 次元ベクトルを

$$\mathbf{w} = [\mathbf{w}[1]^T, \mathbf{w}[2]^T, \dots, \mathbf{w}[F]^T]^T \quad (16)$$

$$\mathbf{w}[f] = \mathcal{V}(W[f]) \quad (17)$$

として、 $\mathbf{w}[f] = \mathcal{V}(\mathcal{M}(\mathbf{w}[f])) = \mathcal{V}(\mathcal{M}(\mathbf{w}))[f]$ のように書くことにすると、式 (14) は

$$\underset{\mathbf{w}}{\text{Minimize}} \quad \mathcal{P}(X\mathbf{w}) - \sum_{f=1}^F \sum_{m=1}^M \log \sigma_m(\mathcal{M}(\mathbf{w}))[f] \quad (18)$$

と書き直せる。ただし、 $FMT \times FM^2$ 行列 X は

$$X = \text{blkdiag}(\boldsymbol{\chi}[1], \boldsymbol{\chi}[2], \dots, \boldsymbol{\chi}[F]) \quad (19)$$

$$\boldsymbol{\chi}[f] = \text{blkdiag}(\chi[f], \chi[f], \dots, \chi[f]) \quad (20)$$

$$\chi[f] = [\tau_1[f], \tau_2[f], \dots, \tau_M[f]] \quad (21)$$

$$\tau_m[f] = [x_m[1, f], x_m[2, f], \dots, x_m[T, f]]^T \quad (22)$$

のように構成され、 $\text{blkdiag}(\cdot)$ は入力された行列たちを単一のブロック対角行列にして返す作用素である。ここで、各変数のサイズは、 $\tau_m[f]$ が $T \times 1$ 、 $\chi[f]$ が $T \times M$ 、 $\boldsymbol{\chi}[f]$ が $MT \times M^2$ である。

簡単のために式 (18) の第二項を

$$\mathcal{I}(\mathbf{w}) = - \sum_{f=1}^F \sum_{m=1}^M \log \sigma_m(\mathcal{M}(\mathbf{w}))[f] \quad (23)$$

と書くことにすれば、式 (18) は

$$\underset{\mathbf{w}}{\text{Minimize}} \quad \mathcal{I}(\mathbf{w}) + \mathcal{P}(X\mathbf{w}) \quad (24)$$

と書ける。これは、主双対分離アルゴリズムを適用可能な式 (8) と同様の形式をしていることがわかる。

¹式 (3) と異なり、式 (14) は分離フィルタ $W[f]$ が正方形でない場合 ($N \neq M$) にも定義可能である。これは、例えば過決定 ($M > N$) の場合に主成分分析などによる次元圧縮を必要としないことを意味し、文献 [1] で述べられているような次元圧縮の悪影響を、本提案手法では回避できる可能性を示唆している。

Algorithm 1 PDS-BSS

```
1: Input:  $X, \mathbf{w}^{[1]}, \mathbf{y}^{[1]}, \mu_1, \mu_2, \alpha$   
2: Output:  $\mathbf{w}^{[K+1]}$   
3: for  $k = 1, \dots, K$  do  
4:    $\tilde{\mathbf{w}} = \text{prox}_{\mu_1 \mathcal{I}}[\mathbf{w}^{[k]} - \mu_1 \mu_2 X^H \mathbf{y}^{[k]}]$   
5:    $\mathbf{z} = \mathbf{y}^{[k]} + X(2\tilde{\mathbf{w}} - \mathbf{w}^{[k]})$   
6:    $\tilde{\mathbf{y}} = \mathbf{z} - \text{prox}_{\frac{1}{\mu_2} \mathcal{P}}[\mathbf{z}]$   
7:    $\mathbf{y}^{[k+1]} = \alpha \tilde{\mathbf{y}} + (1 - \alpha) \mathbf{y}^{[k]}$   
8:    $\mathbf{w}^{[k+1]} = \alpha \tilde{\mathbf{w}} + (1 - \alpha) \mathbf{w}^{[k]}$   
9: end for
```

3.3 提案アルゴリズム

式 (24) に主双対分離アルゴリズムを適用するには、近接作用素を計算する必要がある。式 (23) は行列の特異値に対する関数になっており、そのような関数の近接作用素は、特異値に対する近接作用素で書けることが知られている [6]。その結果、 $-\log \sigma_m$ を $-\log |\sigma_m|$ と思えば、 $\mathcal{I}(\mathbf{w})$ の近接作用素は

$$(\text{prox}_{\mu \mathcal{I}}[\mathbf{w}])(f) = \mathcal{V}(U \tilde{\Sigma}(\mathcal{M}(\mathbf{w}))(f) V^H) \quad (25)$$

と与えられる。ただし、 $W = U \Sigma V^H$ は W の特異値分解、対角行列 $\tilde{\Sigma}(W)$ は

$$\text{diag}(\text{prox}_{-\mu \log}[\sigma_1(W)], \dots, \text{prox}_{-\mu \log}[\sigma_M(W)]) \quad (26)$$

であり、 $\text{prox}_{-\mu \log}[\cdot]$ は式 (11) に示してある。すなわち、 $-\mu \log$ の近接作用素を $W[f]$ の各特異値に適用すれば、 $\text{prox}_{\mu \mathcal{I}}[\cdot]$ を計算することができる。式 (11) を考えれば、この操作は $\|\mathbf{w}\|_2$ を大きく拡大する心配がないという点で、補助関数法に現れがちな逆行行列計算よりも安定して計算可能であると言える。

式 (25) を用いて主双対分離アルゴリズムを適用すれば、**Algorithm 1** を得る。このアルゴリズムは、6 行目に現れる \mathcal{P} の近接作用素 $\text{prox}_{\mathcal{P}/\mu_2}[\cdot]$ を変更するのみで、多様な音源分離モデルに適用することが可能である。従って、式 (12) や (13) のように近接作用素が既知であれば、それらを音源項とするモデルに合わせてアルゴリズムを修正することは非常に容易である。また、 \mathcal{P} の近接作用素が解析的に書けない場合にも、式 (10) の最適化問題を反復最適化することで提案アルゴリズムを適用することが可能である。

3.4 音源項が複数ある問題に対するアルゴリズム

音源項が複数存在する音源分離モデル

$$\underset{\mathbf{w}}{\text{Minimize}} \quad \mathcal{I}(\mathbf{w}) + \sum_{q=1}^Q \mathcal{P}_q(X\mathbf{w}) \quad (27)$$

にも、 $L = [X^T, \dots, X^T]^T$ と思うことで提案手法を適用可能である。式 (27) に対する **Algorithm 1** の拡張を **Algorithm 2** に示す。各音源項 \mathcal{P}_q がそれぞれの近接作用素で独立に扱われているので、音源項が複雑な場合にも、近接作用素が計算可能な関数の和に分

Algorithm 2 PDS-BSS-multiPenalty

```
1: Input:  $X, \mathbf{w}^{[1]}, \mathbf{y}_1^{[1]}, \dots, \mathbf{y}_Q^{[1]}, \mu_1, \mu_2, \alpha$   
2: Output:  $\mathbf{w}^{[K+1]}$   
3: for  $k = 1, \dots, K$  do  
4:    $\tilde{\mathbf{w}} = \text{prox}_{\mu_1 \mathcal{I}}[\mathbf{w}^{[k]} - \mu_1 \mu_2 X^H (\sum_{q=1}^Q \mathbf{y}_q^{[k]})]$   
5:   for  $q = 1, \dots, Q$  do  
6:      $\mathbf{z}_q = \mathbf{y}_q^{[k]} + X(2\tilde{\mathbf{w}} - \mathbf{w}^{[k]})$   
7:      $\tilde{\mathbf{y}}_q = \mathbf{z}_q - \text{prox}_{\frac{1}{\mu_2} \mathcal{P}_q}[\mathbf{z}_q]$   
8:      $\mathbf{y}_q^{[k+1]} = \alpha \tilde{\mathbf{y}}_q + (1 - \alpha) \mathbf{y}_q^{[k]}$   
9:   end for  
10:   $\mathbf{w}^{[k+1]} = \alpha \tilde{\mathbf{w}} + (1 - \alpha) \mathbf{w}^{[k]}$   
11: end for
```

解できれば、**Algorithm 2** によって容易に対応可能である。そのように単純な項の和で表せる音源モデルとしては、例えばロバスト主成分分析 (RPCA) [10] や調波打楽器音分離 (HPSS) [11] などが考えられる。

3.5 ステップサイズパラメータ選択のための正規化

式 (9) の主双対分離アルゴリズムが適切に動作するためには、行列 L のスペクトルノルム $\|\cdot\|_s$ に対して $\mu_1 \mu_2 \|L\|_s^2 \leq 1$ となるようにステップサイズを決定すれば良いことが知られている [7]。これは、 $\|L\|_s = 1$ の場合に $\mu_1 = \mu_2 = 1$ とできることを意味している。従って、観測信号から構成される行列 X を

$$\tilde{X} = X / (\sqrt{Q} \|X\|_s) \quad (28)$$

と正規化することを提案する。すると、ステップサイズがどちらも 1 で良くなるのみでなく、2 章の定数 C と同様の正規化効果が期待できるので、対応する統計モデルから正規化定数 C を計算するのが困難な \mathcal{P} を用いた場合でも、統一的な扱いが可能となる。速度調節のパラメータ $0 < \alpha < 2$ は自由に選ぶことができるので [7]、この正規化によって収束に関わる全てのパラメータを容易に選択できるようになる。

4 実験

IVA と 3 つの新たな分離モデルに提案手法を適用し、その有効性を確認した。音源には SiSEC の一部である UND タスク内の dev1 に含まれる liverec の中で、残響時間 130 ms の環境でマイクロホン間隔 5 cm で収録された女声の音源を用い、 -50° と 45° から到来した音源の混合を **Mixture A**、 -10° と 15° からの混合を **Mixture B** とした。128 ms の Hann 窓をずらし幅 64 ms で用い、各パラメータは $\mu_1 = 1$ 、 $\mu_2 = 1$ 、 $\alpha = 1.75$ とした。また、全ての $W[f]$ は単位行列で初期化し、 \mathbf{y} の初期値はゼロベクトルとした。

4.1 補助関数 IVA (auxIVA) との比較

補助関数法 [4] との比較を行うために、式 (5) に示した球対称 Laplace 分布に基づく IVA

$$\underset{\mathbf{w}}{\text{Minimize}} \quad \mathcal{I}(\mathbf{w}) + \|\tilde{X}\mathbf{w}\|_{2,1} \quad (29)$$

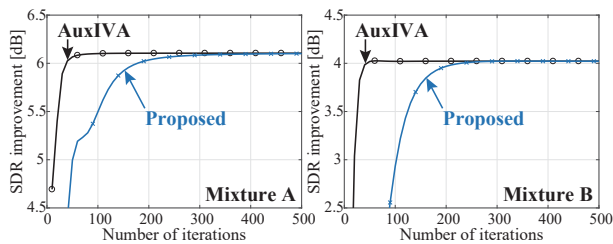


図-1 IVA に対する補助関数法 [4] と提案法の比較。

を考える。図-1 に、補助関数法 [4] と、提案法 (Algorithm 1 に式 (13) の近接作用素を組み込んだもの) による、各反復における SDR の改善量を示す。補助関数法は少ない反復回数で特定の SDR に到達したが、提案法もある程度の反復の後に同様の SDR に到達しており、提案アルゴリズムが適切に機能していることを示唆している。Core i5-7200U による反復毎の平均計算時間は、補助関数法が 80.8 ms だったのに対し、提案法は 47.4 ms であり、計算時間で見ると両者の差は図-1 より小さい。また、補助関数法では、反復回数が多いと破綻する音源も存在し、何らかの安定化を行う必要があったが、提案法はそのような対策をせずとも安定的に実行可能であったことを付記する。

4.2 様々な音源モデルへの応用

提案法の柔軟性を示すために、3つの新たな分離モデルを考え、IVA と比較を行った。1つ目は、スペクトログラムの低ランク性に鑑み [1, 2], 低ランク性を誘導する核ノルム $\|\cdot\|_*$ を用いたモデル (**low-rank**)

$$\underset{\mathbf{w}}{\text{Minimize}} \mathcal{I}(\mathbf{w}) + \|\mathcal{M}_*(\tilde{\mathbf{X}}\mathbf{w})\|_* \quad (30)$$

を考える。ここで、 \mathcal{M}_* は分離後の音源 $\tilde{\mathbf{X}}\mathbf{w}$ を音源毎に行列化する作用素であり、核ノルムは各行列 (スペクトログラム) 毎に適用する。核ノルムの近接作用素はよく知られている通り、特異値に対する soft-thresholding であり [6], 解析的に計算可能である。更に、スパース性を促進することで分離結果が改善されることを期待し、スパース性を誘導する l_1 ノルムを正則化項として加えた IVA (**sparse IVA**)

$$\underset{\mathbf{w}}{\text{Minimize}} \mathcal{I}(\mathbf{w}) + \|\tilde{\mathbf{X}}\mathbf{w}\|_{2,1} + \lambda \|\tilde{\mathbf{X}}\mathbf{w}\|_1 \quad (31)$$

と、スパースな低ランクモデル (**sparse & low-rank**)

$$\underset{\mathbf{w}}{\text{Minimize}} \mathcal{I}(\mathbf{w}) + \|\mathcal{M}_*(\tilde{\mathbf{X}}\mathbf{w})\|_* + \lambda \|\tilde{\mathbf{X}}\mathbf{w}\|_1 \quad (32)$$

も考え、比較を行った。ただし、 $\lambda > 0$ は正則化パラメータを表し、ここでは $\lambda = 0.002$ とした。

図-2 に、提案法による各反復での SDR の改善量を示す。ただし、図-1 の結果も一緒に表示した。スパース正則化を行ったモデルが IVA よりも SDR を改善できており、そのように複数の音源項をもつ分

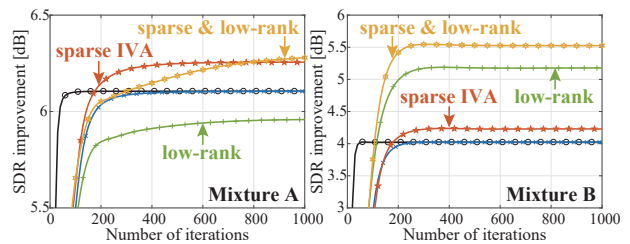


図-2 各モデルによる分離結果 (青線と黒線は図-1 と同じ)。

離モデルの有用性が示唆される。IVA を含めたこれら 4つの音源分離モデルに提案手法を適用するにあたり、ソースコードの修正に要した時間は MATLAB で数分であり、提案法の柔軟さと簡単さが伺える。

5 むすび

独立性に基づく音源分離問題を解くために、主双対分離アルゴリズムを提案した [9]。提案手法では、音源項の近接作用素さえ計算できれば、複数の音源項からなる複雑なモデルも簡単に扱うことが可能である。本提案により、多様なモデルの有効性を統一的にテストするフレームワークが実現したと言える。これを利用し、新たな音源分離モデルの発見が期待される。

参考文献

- [1] D. Kitamura, N. Ono, H. Sawada, H. Kameoka and H. Saruwatari, "Relaxation of rank-1 spatial constraint in overdetermined blind source separation," *Proc. Eur. Signal Process. Conf.*, pp. 1261–1265 (2015).
- [2] D. Kitamura, N. Ono, H. Sawada, H. Kameoka and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, **24**(9), 1626–1641 (2016).
- [3] N. Ono and S. Miyabe, "Auxiliary-function-based independent component analysis for super-gaussian sources," *Proc. Int. Conf. Latent Variable Anal. Signal Separation*, pp. 165–172 (2010).
- [4] N. Ono, "Stable and fast update rules for independent vector analysis based on auxiliary function technique," *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, pp. 189–192 (2011).
- [5] P. L. Combettes and J.-C. Pesquet, *Proximal Splitting Methods in Signal Processing* (Springer, New York, 2011), pp. 185–212.
- [6] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, **1**(3), 127–239 (2014).
- [7] N. Komodakis and J. C. Pesquet, "Playing with duality: An overview of recent primal-dual approaches for solving large-scale optimization problems," *IEEE Signal Process. Mag.*, **32**(6), 31–54 (2015).
- [8] M. Burger, A. Sawatzky and G. Steidl, *First Order Algorithms in Variational Image Processing* (Springer, Cham, 2016), pp. 345–407.
- [9] K. Yatabe and D. Kitamura, "Determined blind source separation via proximal splitting algorithm," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, (submitted).
- [10] Y. Ikemiya, K. Itoyama and K. Yoshii, "Singing voice separation and vocal F0 estimation based on mutual combination of robust principal component analysis and sub-harmonic summation," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, **24**(11), 2084–2095 (2016).
- [11] N. Ono, K. Miyamoto, J. L. Roux, H. Kameoka and S. Sagayama, "Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram," *Proc. Eur. Signal Process. Conf.*, (2008).