

# コサイン類似度罰則条件付き半教師あり非負値行列因子分解\*

☆岩瀬佑太, 北村大地 (香川高専)

## 1 はじめに

音源分離とは、複数の音源が混合した状態で観測された信号から、混合前の個々の音源を分離・抽出する技術である。特に、非負値行列因子分解 (nonnegative matrix factorization: NMF) [1, 2] を用いた音源分離手法は、音源分離の条件や用途等に応じてこれまで数多く提案されてきた [3]–[6]。NMF は Fig. 1 に示すように、観測された非負の時間周波数行列 (振幅スペクトログラム等) を別の二つの非負行列 (基底行列及びアクティベーション行列) の行列積で近似するアルゴリズムである。基底行列には、観測行列中の頻出スペクトルパターンが基底ベクトルとして含まれ、それらの時間的な強度変化がアクティベーション行列に含まれる。これらの特徴量を上手く活用することで、音源分離ができる。

近年では深層学習による音源分離手法が多くのタスクにおいて高い精度を達成しているが、分離対象となる音源 (目的音源) の学習用データセットが極僅かしか用意できないような状況においては、省サンプルの学習データから特徴量を学習しそのまま活用する教師あり・半教師あり NMF [7, 8] が依然として強力である。本稿では、半教師あり NMF (semi-supervised NMF: SNMF) に基づく音源分離のみを取り扱う。

Fig. 2 は SNMF における音源分離アルゴリズムを示しており、学習ステージと分離ステージから構成される。学習ステージでは、目的音源のサンプル信号のスペクトログラム  $\mathbf{X}$  を NMF で分解し、頻出スペクトルパターンを含む教師基底行列  $\mathbf{F}$  を学習する。分離ステージでは、事前学習した教師基底行列を用いて混合信号のスペクトログラム  $\mathbf{Y}$  を分解することで、目的音源成分  $\mathbf{FG}$  とその他の音源 (非目的音源) 成分  $\mathbf{HU}$  を推定する。

SNMF に基づく音源分離では、混合信号中の目的音源と非目的音源が類似したスペクトルを含む場合、分離精度が劣化する問題がある。これは、類似するスペクトルが教師基底行列  $\mathbf{F}$  又は非目的音源の基底行列  $\mathbf{H}$  のいずれを用いても表現できることが原因である。この場合、目的音源の一部が  $\mathbf{HU}$  に、あるいは非目的音源の一部が  $\mathbf{FG}$  に混入することとなり、音源分離精度が劣化する。

この問題に対処するために、罰則条件付き SNMF (penalized SNMF: PSNMF) [8] が提案されている。PSNMF は、教師基底行列  $\mathbf{F}$  と非目的音源の基底行列  $\mathbf{H}$  ができるだけ直交化する罰則条件を付与したものであり、目的音源成分  $\mathbf{FG}$  の品質を大幅に向上させる効果がある。しかし、既存の PSNMF は、NMF が持つ基底行列・アクティベーション行列間のスケール不定性により、罰則条件が正しく影響しない定式化となっている。同様の現象は、アクティベーション行列にスパース正則化を与える NMF でも確認されており、正しい制約条件を与えた最適化法が提案されている [9]。

本稿では、PSNMF における直交化罰則条件をス

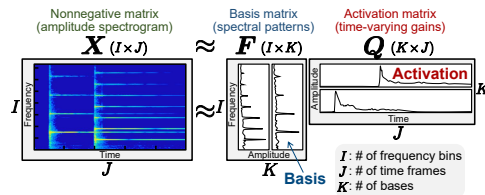


Fig. 1: Matrix decomposition by NMF.

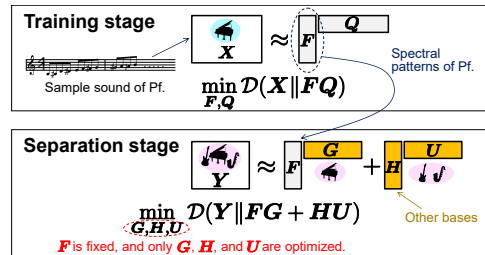


Fig. 2: Separation algorithm in SNMF.

ケール不定性に依らない形で定式化し最適化する手法を新たに提案する。提案手法では、コサイン類似度を用いた罰則条件を SNMF に付与し、majorization-minimization (MM) アルゴリズム [10] による乗法型更新則を導出する。また、音楽信号を対象とした音源分離に適用し、従来手法よりも優れた音源分離が達成できることを実験的に示す。

## 2 従来手法

### 2.1 NMF 及び SNMF

NMF の最適化問題は次式で表される。

$$\min_{\mathbf{F}, \mathbf{Q}} \mathcal{D}(\mathbf{X} \| \mathbf{F}\mathbf{Q}) \text{ s.t. } f_{ik}, q_{kj} \geq 0 \forall i, j, k \quad (1)$$

ここで、 $\mathbf{X} \in \mathbb{R}_{\geq 0}^{I \times J}$  は非負観測行列であり、本稿では振幅スペクトログラムとする。また、 $\mathbf{F} \in \mathbb{R}_{\geq 0}^{I \times K}$  及び  $\mathbf{Q} \in \mathbb{R}_{\geq 0}^{K \times J}$  はそれぞれ基底行列及びアクティベーション行列であり、 $f_{ik}$  及び  $q_{kj}$  はそれぞれ  $\mathbf{F}$  及び  $\mathbf{Q}$  の要素である。さらに、 $i = 1, \dots, I$ ,  $j = 1, \dots, J$ , 及び  $k = 1, \dots, K$  はそれぞれ周波数ビン、時間フレーム、及び基底ベクトルのインデックスを表す。 $\mathcal{D}(\mathbf{A} \| \mathbf{B})$  は、行列  $\mathbf{A} \in \mathbb{R}_{\geq 0}^{I \times J}$  及び  $\mathbf{B} \in \mathbb{R}_{\geq 0}^{I \times J}$  間の乖離度 (ダイバージェンス) を返す関数である。本稿では、次式に示す一般化カルバック・ライブラダイバージェンスを用いる。

$$\mathcal{D}(\mathbf{A} \| \mathbf{B}) = \sum_{i,j} \left( a_{ij} \log \frac{a_{ij}}{b_{ij}} - a_{ij} + b_{ij} \right) \quad (2)$$

ここで、 $a_{ij}$  及び  $b_{ij}$  はそれぞれ行列  $\mathbf{A}$  及び  $\mathbf{B}$  の要素である。従って、式 (1) の最小化問題を解くことで  $\mathbf{F}$  と  $\mathbf{Q}$  が推定される。

\*Semi-supervised nonnegative matrix factorization with cosine similarity penalty by Yuta Iwase (NIT Kagawa) and Daichi Kitamura (NIT Kagawa).

SNMF では、学習ステージにおいて目的音源のサンプル信号  $\mathbf{X}$  に NMF を適用することで、教師基底行列  $\mathbf{F}$  を事前学習する。このとき、 $\mathbf{F}$  は目的音源の頻出スペクトルパターンを  $K$  本の列ベクトル（基底ベクトル）として含んでいる。分離ステージでは、教師基底行列  $\mathbf{F}$  が与えられた下で、混合信号の振幅スペクトログラム  $\mathbf{Y} \in \mathbb{R}_{\geq 0}^{I \times J}$  を次の最適化問題により分解する。

$$\begin{aligned} \min_{\mathbf{G}, \mathbf{H}, \mathbf{U}} \mathcal{D}(\mathbf{Y} \|\mathbf{FG} + \mathbf{HU}) \\ \text{s.t. } g_{kj}, h_{il}, u_{lj} \geq 0 \quad \forall i, j, k, l \end{aligned} \quad (3)$$

ここで、 $\mathbf{H} \in \mathbb{R}_{\geq 0}^{I \times L}$  及び  $\mathbf{U} \in \mathbb{R}_{\geq 0}^{L \times J}$  はそれぞれ非目的音源を表すための基底行列とアクティベーション行列であり、 $h_{il}$  及び  $u_{lj}$  はそれぞれ  $\mathbf{H}$  及び  $\mathbf{U}$  の要素である。また、 $l = 1, \dots, L$  は非目的音源の基底行列  $\mathbf{H}$  に含まれる基底のインデックスを表す。理想的には、混合信号  $\mathbf{Y}$  中の目的音源の成分が  $\mathbf{FG}$ 、非目的音源の成分が  $\mathbf{HU}$  として分離されるため、音源分離が可能となる。しかし、混合信号中の目的音源と非目的音源が類似したスペクトルを含む場合、そのような成分は  $\mathbf{F}$  又は  $\mathbf{H}$  のいずれを用いても表現できるため、目的音源の一部が  $\mathbf{HU}$  に、あるいは非目的音源の一部が  $\mathbf{FG}$  に混入し、音源分離精度が劣化する。

## 2.2 直交化に基づく PSNMF

前述の問題に対処するために、SNMF の最適化に対して、非目的音源の基底行列  $\mathbf{H}$  が教師基底行列  $\mathbf{F}$  とできるだけ直交化する罰則条件を与えた手法 (PSNMF) が提案されている [8]。本手法では、分離ステージにおいて次の最適化問題を考える。

$$\begin{aligned} \min_{\mathbf{G}, \mathbf{H}, \mathbf{U}} \mathcal{D}(\mathbf{Y} \|\mathbf{FG} + \mathbf{HU}) + \mu \mathcal{P}_{\text{orth}}(\mathbf{F}, \mathbf{H}) \\ \text{s.t. } g_{kj}, h_{il}, u_{lj} \geq 0 \quad \forall i, j, k, l \end{aligned} \quad (4)$$

$$\begin{aligned} \mathcal{P}_{\text{orth}}(\mathbf{F}, \mathbf{H}) &= \|\mathbf{F}^T \mathbf{H}\|_{\text{Fr}}^2 \\ &= \sum_{k, l} \left( \sum_i f_{ik} h_{il} \right)^2 \end{aligned} \quad (5)$$

ここで、 $\mu$  は罰則項  $\mathcal{P}_{\text{orth}}(\mathbf{F}, \mathbf{H})$  の重み係数、 $\|\mathbf{A}\|_{\text{Fr}}$  は行列  $\mathbf{A}$  のフロベニウスノルムである。罰則項  $\mathcal{P}_{\text{orth}}(\mathbf{F}, \mathbf{H})$  は、 $\mathbf{F}$  の基底ベクトル ( $K$  本) と  $\mathbf{H}$  の基底ベクトル ( $L$  本) の総組み合わせの内積の二乗和に対応する。従って、 $\mathbf{H}$  を推定する際には、「 $\mathbf{Y}$  と  $\mathbf{FG} + \mathbf{HU}$  が近くなること」及び「 $\mathbf{H}$  の基底ベクトルが  $\mathbf{F}$  の基底ベクトルとなるべく直交すること」の2点が考慮され、結果として SNMF よりも音源分離が促進されることが実験的に確認されている [8]。

MM アルゴリズム [10] に基づく変数行列  $\mathbf{G}$ 、 $\mathbf{H}$ 、及び  $\mathbf{U}$  の更新則はそれぞれ次のようになる [8]。

$$g_{kj} \leftarrow g_{kj} \cdot \frac{\sum_i \frac{y_{ij}}{\sum_{k'} f_{ik'} g_{k'j} + \sum_{l'} h_{il'} u_{l'j}} f_{ik}}{\sum_i f_{ik}} \quad (6)$$

$$h_{il} \leftarrow h_{il} \cdot \frac{\sum_j \frac{y_{ij}}{\sum_{k'} f_{ik'} g_{k'j} + \sum_{l'} h_{il'} u_{l'j}} u_{lj}}{\sum_j u_{lj} + \mu f_{ik} \sum_{i'} f_{i'k} h_{i'l}} \quad (7)$$

$$u_{lj} \leftarrow u_{lj} \cdot \frac{\sum_i \frac{y_{ij}}{\sum_{k'} f_{ik'} g_{k'j} + \sum_{l'} h_{il'} u_{l'j}} h_{il}}{\sum_i h_{il}} \quad (8)$$

ここで、 $y_{ij}$  は  $\mathbf{Y}$  の要素である。

## 3 提案手法

### 3.1 動機

直交化に基づく PSNMF には、罰則項の効果が直接反映されない本質的な問題がある。これは、NMF が本来持っている基底行列・アクティベーション行列間のスケール不定性が原因である。即ち、 $\mathbf{X} \approx \mathbf{FQ}$  という NMF 分解を考えた際に、基底行列  $\mathbf{F}$  に任意の正の係数  $\alpha$  を乗じて  $\alpha \mathbf{F}$  と変形しても、同時にアクティベーション行列  $\mathbf{Q}$  を  $(1/\alpha) \mathbf{Q}$  とすれば、 $\mathbf{FQ}$  自体の値は変わらないという不定性である。

直交化に基づく PSNMF では、罰則項  $\mathcal{P}_{\text{orth}}(\mathbf{F}, \mathbf{H})$  の値を小さくすることで  $\mathbf{F}$  と  $\mathbf{H}$  の非類似性を高めることを考えるが、実際にはこの罰則項は非常に小さな値の  $\alpha$  を乗じて  $\alpha \mathbf{H}$  とすることで最小化可能である。このとき、 $\mathbf{HU}$  の値が変わらないように、 $\mathbf{U}$  は  $(1/\alpha) \mathbf{U}$  と更新される。従って、この罰則項は直接的に  $\mathbf{F}$  と  $\mathbf{H}$  を直交化していない。

この問題を避けるために、文献 [8] では、 $\mathbf{H}$  を更新する度に全基底ベクトルを正規化する（正規化係数の逆数を  $\mathbf{U}$  の行ベクトルに乗じて  $\mathbf{HU}$  は不変とする）ように実装しており、実験的には性能改善が見られたが、この実装方法も根本的な解決とは言えない。同様の現象は、 $L_1$  ノルムに基づくスパース正則化をアクティベーション行列に与える NMF でも確認されており、基底ベクトルの長さを制約条件に与える最適化法が提案されている [9]。PSNMF においても、文献 [9] と同様の制約条件を導入して解くことはできるが、よりシンプルな解決法として、基底ベクトル間の角度にのみ依存するコサイン類似度を新たに罰則項に用いる PSNMF を提案する。

### 3.2 コサイン類似度に基づく PSNMF

コサイン類似度に基づく PSNMF として、次の最適化問題を考える。

$$\begin{aligned} \min_{\mathbf{G}, \mathbf{H}, \mathbf{U}} \mathcal{D}(\mathbf{Y} \|\mathbf{FG} + \mathbf{HU}) + \mu \mathcal{P}_{\text{cos}}(\mathbf{F}, \mathbf{H}) \\ \text{s.t. } g_{kj}, h_{il}, u_{lj} \geq 0 \quad \forall i, j, k, l \end{aligned} \quad (9)$$

$$\mathcal{P}_{\text{cos}}(\mathbf{F}, \mathbf{H}) = \sum_{k, l} \log \frac{\sum_i f_{ik} h_{il}}{\left( \sum_i f_{ik}^2 \right)^{\frac{1}{2}} \left( \sum_i h_{il}^2 \right)^{\frac{1}{2}}} \quad (10)$$

ここで、罰則項  $\mathcal{P}_{\text{cos}}(\mathbf{F}, \mathbf{H})$  は、 $\mathbf{F}$  の基底ベクトルと  $\mathbf{H}$  の基底ベクトルの総組み合わせのコサイン類似度の対数和に対応する。コサイン類似度であるので、基底ベクトルの長さには依らず、角度のみで直交性を測ることができ、NMF におけるスケールの不定性は影響しない。また、基底ベクトルは全て非負であることから、コサイン類似度の値は最小で 0（直交）、最大で 1（平行）となる。さらに、コサイン類似度の対数を取ることで、次節に述べる MM アルゴリズムでの導出が可能となる。

### 3.3 MM アルゴリズムに基づく乗法型更新則の導出

式 (9) を直接最小化することは困難なため、従来の PSNMF と同様に MM アルゴリズムを用いて最小化する。なお、 $\mathbf{G}$  と  $\mathbf{U}$  の偏微分は罰則項  $\mathcal{P}_{\cos}$  によらないため、乗法型更新則は式 (6) 及び (8) と同様である。式 (9) は次のように変形できる。

$$\begin{aligned} & \mathcal{D}(\mathbf{Y} \|\mathbf{F}\mathbf{G} + \mathbf{H}\mathbf{U}) + \mu \mathcal{P}_{\cos}(\mathbf{F}, \mathbf{H}) \\ &= \sum_{i,j} \left[ y_{ij} \log y_{ij} - y_{ij} \log \left( \sum_k f_{ik} g_{kj} + \sum_l h_{il} u_{lj} \right) \right. \\ & \quad \left. - y_{ij} + \sum_k f_{ik} g_{kj} + \sum_l h_{il} u_{lj} \right] \\ & \quad + \mu \sum_{k,l} \left[ \log \sum_i f_{ik} h_{il} - \frac{1}{2} \log \sum_i f_{ik}^2 - \frac{1}{2} \log \sum_i h_{il}^2 \right] \end{aligned} \quad (11)$$

上式の第二項及び第八項にイェンセンの不等式、第六項に接線不等式をそれぞれ適用し、上限関数を次のように設計する。 $\sum_k \alpha_{ijk} = 1$ ,  $\sum_l \beta_{ijl} = 1$ ,  $\sum_i \delta_{kli} = 1$ ,  $\sum_i \zeta_{kli} = 1$ ,  $\gamma_{ij1} + \gamma_{ij2} = 1$  をみたす補助変数  $\alpha_{ijk}$ ,  $\beta_{ijl}$ ,  $\delta_{kli}$ ,  $\zeta_{kli}$ ,  $\gamma_{ij1}$ ,  $\gamma_{ij2} > 0$  を用いると、上限関数を次のように設計できる。

$$\begin{aligned} \mathcal{J} \leq & \sum_{i,j} \left[ y_{ij} \log y_{ij} - y_{ij} \left( \gamma_{ij1} \sum_k \alpha_{ijk} \log \frac{f_{ik} g_{kj}}{\alpha_{ijk}} \right. \right. \\ & \quad \left. \left. + \gamma_{ij2} \sum_l \beta_{ijl} \log \frac{h_{il} u_{lj}}{\beta_{ijl}} - \gamma_{ij1} \log \gamma_{ij1} \right. \right. \\ & \quad \left. \left. - \gamma_{ij2} \log \gamma_{ij2} \right) - y_{ij} + \sum_k f_{ik} g_{kj} + \sum_l h_{il} u_{lj} \right] \\ & \quad + \mu \sum_{k,l} \left[ \frac{1}{\zeta_{kli}} \left( \sum_i f_{ik} h_{il} - \zeta_{kli} \right) + \log \zeta_{kli} \right. \\ & \quad \left. - \frac{1}{2} \log \sum_i f_{ik}^2 - \frac{1}{2} \sum_i \delta_{kli} (2 \log h_{il} - \log \delta_{kli}) \right] \end{aligned} \quad (12)$$

ただし、式 (12) の等号成立条件は以下である。

$$\alpha_{ijk} = \frac{f_{ik} g_{kj}}{\sum_{k'} f_{ik'} g_{k'j}} \quad (13)$$

$$\beta_{ijl} = \frac{h_{il} u_{lj}}{\sum_{l'} h_{il'} u_{l'j}} \quad (14)$$

$$\zeta_{kli} = \sum_i f_{ik} h_{il} \quad (15)$$

$$\delta_{kli} = \frac{h_{il}^2}{\sum_{i'} h_{i'l}^2} \quad (16)$$

$$\gamma_{ij1} = \frac{\sum_k f_{ik} g_{kj}}{\sum_k f_{ik} g_{kj} + \sum_l h_{il} u_{lj}} \quad (17)$$

$$\gamma_{ij2} = \frac{\sum_l h_{il} u_{lj}}{\sum_k f_{ik} g_{kj} + \sum_l h_{il} u_{lj}} \quad (18)$$

$\mathbf{H}$  の更新式は式 (12) の右辺を  $h_{il}$  で偏微分し 0 とおいた式から次のように得られる。

$$\begin{aligned} h_{il} \leftarrow & \frac{\sum_j y_{ij} \gamma_{ij2} \beta_{ijl} + \mu \sum_k \delta_{kli}}{\sum_j u_{lj} + \mu \sum_k \frac{f_{ik}}{\zeta_{kli}}} \\ & = h_{il} \cdot \frac{\sum_j \frac{y_{ij}}{\sum_{k'} f_{ik'} g_{k'j} + \sum_{l'} h_{il'} u_{l'j}} u_{lj} + \mu K \frac{h_{il}}{\sum_{i'} h_{i'l}^2}}{\sum_j u_{lj} + \mu \sum_k \frac{f_{ik}}{\sum_{i'} f_{i'k} h_{i'l}}} \end{aligned} \quad (19)$$

## 4 実験

### 4.1 実験条件

提案手法の有効性を確認するために、音楽信号の音源分離性能を直交化に基づく PSNMF (従来手法) とコサイン類似度に基づく PSNMF (提案手法) に関して比較した。実験では、データセットとして songKitamura [6, 11] の YMH MU-1000 の音源を用いた。音源にはオーボエ (Ob.), トランペット (Tp.), ホルン (Hr.), フルート (Fl.), ヴァイオリン (Vn.), クラリネット (Cl.), ピアノ (Pf.), ハープシコード (Hp.), トロンボーン (Tr.), ファゴット (Fg.), チェロ (Vc.) の 11 種類の 2 オクターブ上昇音階 (学習ステージ用) とメロディパート (分離ステージ用) を用いた。11 種の楽器から 2 つを選択し、それらのメロディーパートを等パワーで混合し混合信号とした。従って、合計 90 曲での評価となった。評価指標には、信号対歪み比 (source-to-distortion ratio: SDR) [12] を用いた。

重み係数  $\mu$  は対数的に 72 段階に分割した値を使用した。各行列の初期値は区間 (0, 1) の一様乱数とした。短時間フーリエ変換の窓長とシフト幅はそれぞれ 92.9ms 及び 46.4ms とした。目的音源と非目的音源の基底ベクトル数はそれぞれ  $K = 27$  及び  $L = 50$  とした。

### 4.2 実験結果

Fig. 3 に曲ごとの分離結果の例を示す。また、Table 1 は、各曲の最大 SDR 値を全 90 曲に関して求めた際の平均値と中央値を示す。これらの結果より、コサイン類似度を罰則項に用いる提案手法が従来手法よりも高い SDR 値を記録していることが分かる。但し、提案手法はスケールの不定性が影響しないことから、重み係数  $\mu$  の設定は従来手法よりもピーキーである様子も確認できる。

## 5 まとめ

本稿では、半教師あり NMF において、スケールの不定性の影響を受けない罰則項としてコサイン類似度を用いたアルゴリズムを新たに提案した。評価実験の結果、従来手法よりも高精度な音源分離が可能であることを示した。

謝辞 本研究の一部は JSPS 科研費 19K20306 の助成を受けたものである。

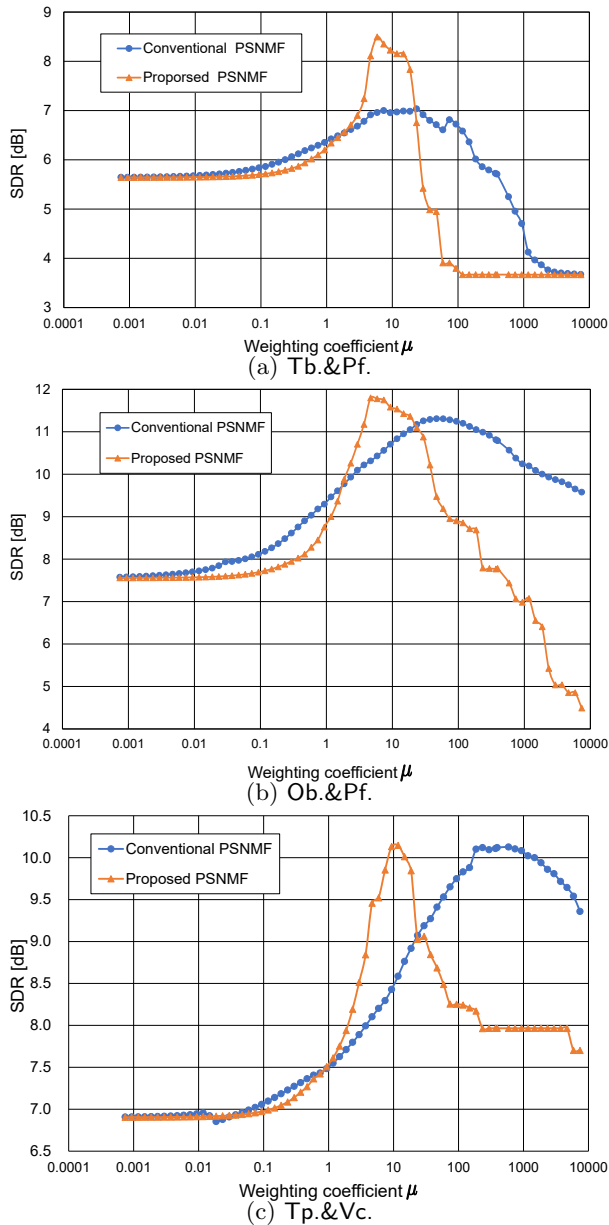


Fig. 3: Examples of SDR values of conventional and proposed PSNMFs for (a) Tb.&Pf., (b) Ob.&Pf., and (c) Tp.&Vc. mixtures.

Table 1: Average and median SDRs for each method

Method	Average	Median
SNMF	6.09 dB	5.73 dB
Conventional PSNMF	8.51 dB	7.91 dB
Proposed PSNMF	8.73 dB	8.22 dB

## 参考文献

- [1] D. D. Lee and H. S. Seung, “Learning the parts of objects by non-negative matrix factorization,” *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [2] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” *Proc. NIPS*, pp. 556–562, 2000.
- [3] T. Virtanen, “Monaural sound source separation by nonnegative matrix factorization

with temporal continuity and sparseness criteria,” *IEEE Trans. ASLP*, vol. 15, no. 3, pp. 1066–1074, 2007.

- [4] A. Ozerov, C. Févotte, and M. Charbit, “Factorial scaled hidden Markov model for polyphonic audio representation and source separation,” *Proc. WASPAA*, pp. 121–124, 2009.
- [5] H. Kameoka, M. Nakano, K. Ochiai, Y. Imoto, K. Kashino, and S. Sagayama, “Constrained and regularized variants of non-negative matrix factorization incorporating music-specific constraints,” *Proc. ICASSP*, pp. 5365–5368, 2012.
- [6] D. Kitamura, H. Saruwatari, H. Kameoka, Y. Takahashi, K. Kondo, and S. Nakamura, “Multichannel signal separation combining directional clustering and nonnegative matrix factorization with spectrogram restoration,” *IEEE/ACM Trans. ASLP*, vol. 23, no. 4, pp. 654–669, 2015.
- [7] P. Smaragdis, B. Raj, and M. Shashanka, “Supervised and semi-supervised separation of sounds from single-channel mixtures,” *Proc. LVA/ICA*, pp. 414–421, 2007.
- [8] D. Kitamura, H. Saruwatari, K. Yagi, K. Shikano, Y. Takahashi, and K. Kondo, “Music signal separation based on supervised non-negative matrix factorization with orthogonality and maximum-divergence penalties,” *IEICE Trans. Fundamentals*, vol. E97-A, no. 5, pp. 1113–1118, 2014.
- [9] J. Le Roux, F. Weninger, and J. R. Hershey, “Sparse NMF - half-baked or well done?,” *Mitsubishi Electric Research Lab. Technical Report*, TR2015-023, 2015.
- [10] D. R. Hunter and K. Lange, “Quantile regression via an MM algorithm,” *J. Comput. Graph. Stat.*, vol. 9, no. 1, pp. 60–77, 2000.
- [11] D. Kitamura, “Open dataset: songKitamura” <http://d-kitamura.net/dataset.html>. Accessed 24 December 2019.
- [12] E. Vincent, R. Gribonval, and C. Févotte, “Performance measurement in blind audio source separation,” *IEEE Trans. ASLP*, vol. 14, no. 4, pp. 1462–1469, 2006.