

## 全極モデルを用いた基底変形型教師あり NMF による音楽信号分離\*

☆中嶋広明 (東大), 北村大地 (総研大), 高宗典玄 (東大), 小山翔一 (東大), 猿渡洋 (東大),  
小野順貴 (NII/総研大), 高橋祐 (ヤマハ), 近藤多伸 (ヤマハ)

## 1 はじめに

近年、低ランク近似手法である非負値行列因子分解 (nonnegative matrix factorization: NMF) による音源分離技術が活発に研究されている。NMF とは、スペクトログラムを非負の基底行列とアクティベーション行列の積に分解する手法である。特に、モノラル・フォーマットで収録・再現される音楽信号の分離において、NMF は有力な手法の一つになっている [1]。

NMF による音源分離手法は、大きく分けて教師なしと教師ありに分けられる。教師なし NMF では、訓練データを与えないまま NMF により音源分離を達成する [2]。教師あり NMF (supervised NMF: SNMF) では、目的音のサンプルを訓練データとして利用することで音源分離を達成する [3]。しかし教師あり NMF において、事前に学習する教師音と実際に分離したい目的音の音色が異なる場合に分離の性能が著しく劣化するという問題点がある。

この問題を解決するために、混合音中に含まれる目的音へ適応して教師基底を変形する NMF を用いた分離アルゴリズムを提案する。提案手法では、線形時不変な全極モデルに基づいて教師基底を変形し、分離を行う。音楽信号分離実験では、提案手法が従来の教師あり NMF よりもよい分離音を得ることができる場合が存在することを確認した。

## 2 従来手法

## 2.1 教師あり NMF 音源分離の概要

教師あり NMF は事前学習と観測混合音分離の 2 段階から構成される。以降で詳細を述べる。

まず、事前学習によって教師基底を得る。式 (1) のように、抽出したい目的音による単独の演奏音を教師音として NMF により基底行列とアクティベーション行列に分解することで、当該目的音に関する教師基底を得る。

$$\mathbf{Y}_{\text{teacher}} \approx \mathbf{F}\mathbf{G} \quad (1)$$

$\mathbf{Y}_{\text{teacher}}$  は  $\Omega$  行  $T_1$  列の非負値行列で、ここでは振幅スペクトログラムを表すものとする。 $\mathbf{F}$  は  $\Omega$  行  $K$  列の非負値行列で一般的に基底行列と呼ばれる。 $\mathbf{G}$  は  $K$  行  $T$  列の非負値行列で、 $\mathbf{F}$  の各基底ベクトルの励起情報を持っていることからアクティベーション行列と呼ばれる。ここで、 $K$  は行列  $\mathbf{F}$  の基底数である。

SNMF の観測混合音の分離は、事前学習で得た基底行列  $\mathbf{F}$  を用いて次式のように書ける。

$$\mathbf{Y}_{\text{mix}} \approx \mathbf{F}\mathbf{G} + \mathbf{H}\mathbf{U} \quad (2)$$

$\mathbf{Y}_{\text{mix}}$  は  $\Omega$  行  $T_2$  列の非負値行列で、ここでは振幅スペクトログラムである。この分離は事前学習で得た基底

行列  $\mathbf{F}$  を与えた上で求めているため、理想的には  $\mathbf{F}\mathbf{G}$  は観測混合音中の抽出したい目的音成分を表し、 $\mathbf{H}\mathbf{U}$  は妨害音を表す。

式 (2) の分解は、以下の最小化問題を解くことに帰着する。

$$\min_{\mathbf{G}, \mathbf{H}, \mathbf{U}} \mathcal{D}_{\text{KL}}(\mathbf{Y}|\mathbf{F}\mathbf{G} + \mathbf{H}\mathbf{U}) \quad (3)$$

ただし  $\mathcal{D}_{\text{KL}}$  は一般化 KL ダイバージェンスであり、次のように定義する。

$$\mathcal{D}_{\text{KL}}(y||x) = y \log \frac{y}{x} + x - y \quad (4)$$

式 (3) の最小化問題は、補助関数法により求めることができる。

## 2.2 加法型基底変形 NMF

2.1 節で述べた SNMF では、事前学習で用いる教師音と観測混合音中の目的音の間で音色が異なる場合に、大きく分離性能が低下するという欠点がある。そこで、目的音に対し事前学習で得た基底行列を変形させることで分離性能の向上を目指す手法として、加法型基底変形 NMF[4] が提案された。

加法型基底変形 NMF における観測混合音の分離は、事前学習で得た基底行列  $\mathbf{F}$  を用いて次式のように書くことができる。

$$\mathbf{Y}_{\text{mix}} \approx (\mathbf{F} + \mathbf{D})\mathbf{G} + \mathbf{H}\mathbf{U} \quad (5)$$

$\mathbf{D}$  は  $\mathbf{F}$  とアクティベーション行列  $\mathbf{G}$  を共有する  $\Omega$  行  $K$  列の基底行列である。この分解では、 $\mathbf{F}$  では表現しきれない目的音の基底を  $\mathbf{D}$  で表現している。 $\mathbf{D}$  は非負値行列ではないが、 $\mathbf{F} + \mathbf{D}$  が非負値になるよう適当な制約を加えるものとする。

しかしこの手法には、3 つ以上の重みパラメータがあるため調整が困難なこと、各変数の初期値依存性が極めて強いこと、線形時不変な変形ではないなどの問題点が存在する。

## 2.3 統計的音声スペクトル推定器を用いたポストフィルタ

一般化 minimum mean-square error short-time spectral amplitude (MMSE-STSA) 推定器 [5] は、目的音を表す事前分布の元で、真の目的音振幅スペクトルとその推定値との平均二乗誤差を最小化するスペクトルゲインを求める。このスペクトルゲインを観測信号スペクトルに乗算することで推定目的音が得られ、音源分離が達成できる。一般化 MMSE-STSA 推定器における推定目的音を  $\mathbf{Y}$  とすると、観測音  $\mathbf{Y}_{\text{mix}}$  から推定目的音  $\mathbf{Y}$  の推定式は式 (6) のように、スペクトルゲイン  $\mathbf{J}$  は式 (7) のように与えられる。ただし  $\circ$  はアダマール積、 $J_{\omega,t}$  は  $\mathbf{J}$  の要素値である。

\*“Music Signal Separation Using Supervised NMF with All-Pole-Model-Based Basis Deformation,” by Hiroaki Nakajima (The University of Tokyo), Daichi Kitamura (SOKENDAI), Norihiro Takamune (The University of Tokyo), Shoichi Koyama (The University of Tokyo), Hiroshi Saruwatari (The University of Tokyo), Nobutaka Ono (NII/SOKENDAI), Yu Takahashi (Yamaha), Kazunobu Kondo (Yamaha).

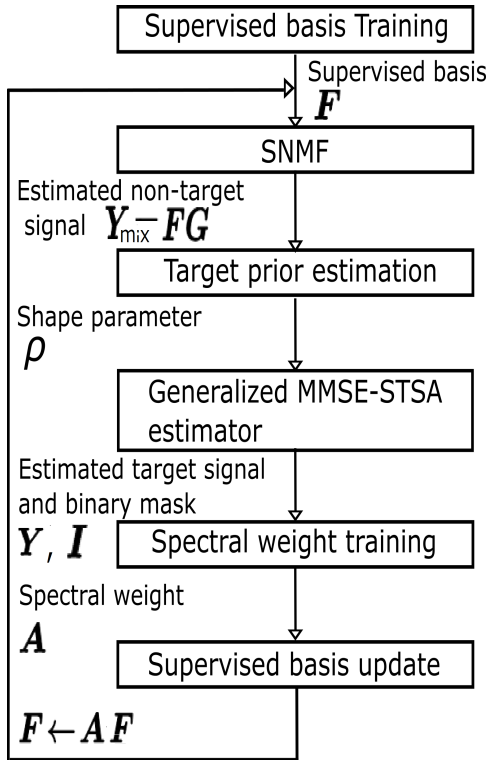


Fig. 1 提案手法のブロック図

$$Y = J \circ Y_{\text{mix}} \quad (6)$$

$$J_{\omega,t} = \frac{\sqrt{v_{\omega,t}}}{\tilde{\gamma}_{\omega,t}} \cdot \left( \frac{\Gamma(\rho + 0.5)}{\Gamma(\rho)} \cdot \frac{\Phi(0.5 - \rho, 1, -v_{\omega,t})}{\Phi(1 - \rho, 1, -v_{\omega,t})} \right)^{1/\beta} \quad (7)$$

ここで  $\Gamma(\cdot)$  はガンマ関数、 $\Phi(a, b; k) = F_1(a, b; k)$  は合流型超幾何関数、 $\beta$  は振幅圧縮パラメータ、 $\rho$  は事前分布であるカイ分布における形状母数を表す。また、 $v_{\omega,t}$  は事前 SN 比  $\tilde{\epsilon}_{\omega,t}$ 、事後 SN 比  $\tilde{\gamma}_{\omega,t}$  を用いて次のように与えられる。

$$v_{\omega,t} = \tilde{\gamma}_{\omega,t} \tilde{\epsilon}_{\omega,t} (1 + \tilde{\epsilon}_{\omega,t})^{-1} \quad (8)$$

一般化 MMSE-STSA 推定器では、 $\tilde{\gamma}_{\omega,t}$  の計算に妨害音のパワースペクトルが必要になるが、本研究では基底変形前の教師基底による SNMF によって得られる妨害音  $Y_{\text{mix}} - FG$  を用いる。

### 3 提案手法

#### 3.1 提案手法の概要

教師音と目的音の音色が異なる際に分離性能が著しく低下する原因として、事前学習で得た教師基底が目的音の分離に適していないことが考えられる。そこで、教師基底を変形する必要がある。しかし従来の基底変形型 SNMF では、教師基底の変形と信号分離を同時に行うため、最適化が困難である。本研究では、SNMF において一般化 MMSE-STSA 推定器の出力値を用いることで信号分離とは別に教師基底変形を行う手法を提案する。

本稿で提案する処理を Fig. 1 に示す。本手法では、観測音信号に対する周波数方向への適応を SNMF に

おける教師基底変形により行い、時間方向への適応を目的音の事前分布パラメータ推定により行う。具体的には、第一に現在得ている教師基底を用いて SNMF を行う。得られた妨害音を初期値として一般化 MMSE-STSA 推定器を適用し、時間周波数領域で推定目的音及び各成分の確信度をしめすバイナリマスクが得られる。最後に以下で提案する全極モデルによる教師基底の変換を施し、変形された教師基底を新しい教師基底とする。これらの処理を任意の反復回数だけ行い、得られる新しい教師基底による SNMF の推定目的音を最終的な出力とする。

#### 3.2 一般化 MMSE-STSA 推定器の出力を用いた全極モデルによる教師基底の変形

安良岡ら [6] は音響信号の音色構造因子を推定するため、スペクトル包絡を全極モデルで表現する NMF を提案した。これを参考にして、本節では一般化 MMSE-STSA 推定器の出力を用いた全極モデルによる教師基底の変形手法について述べる。提案手法では、教師音を用いて教師基底  $F$  を学習し、その後で前の反復により得られた推定目的音を用いて教師基底の変形成分を学習する。教師基底の変形成分の学習は次のように行う。

$$I \circ Y \approx I \circ (AFG) \quad (9)$$

ここで  $I$  は各成分の確信度を 1 と 0 で示すバイナリマスク、 $A$  は対角成分に全極モデルによるスペクトル重みを持つ行列を表す。 $A$  の対角成分は、次のように全極モデルで与える。

$$A_{\omega,\omega} = \frac{1}{|1 - \sum_{k=1}^p \alpha_k \exp(-\pi j k \frac{\omega}{\Omega})|} \quad (10)$$

ここで  $p$  は次数、 $\alpha_k$  は全極モデルの係数である。また、以後の計算の便宜のため  $A_{\omega} = 1 - \sum_{k=1}^p \alpha_k \exp(-\pi j k \frac{\omega}{\Omega})$  と置く。

この学習では、周波数ごとに全極モデルで表現されるスペクトル重み  $A$  を学習し、教師基底に乗ずることで観測音信号中の目的音信号に適応させる。これにより、教師音と目的音の音色の違いを表現することを期待している。

#### 3.3 目的関数と更新式

本節では、スペクトル重みの学習に用いる NMF の目的関数および更新式の導出過程を示す。本研究では、一般化 KL ダイバージェンスを NMF の目的関数として用いる。一般化 KL ダイバージェンスによる全極モデルを用いた基底変形型教師あり NMF の目的関数  $\mathcal{J}$  は次のように表される。

$$\mathcal{J} = \sum_{\omega,t} i_{\omega,t} \left\{ -y_{\omega,t} + \frac{\sum_k f_{\omega,k} g_{k,t}}{|A_{\omega}|} + y_{\omega,t} \log \frac{y_{\omega,t}}{\sum_k f_{\omega,k} g_{k,t} / |A_{\omega}|} \right\} \quad (11)$$

ここで  $i_{\omega,t}$ 、 $y_{\omega,t}$ 、 $f_{\omega,k}$  及び  $g_{k,t}$  はそれぞれ  $I$ 、 $Y$ 、 $F$ 、 $G$  に対応する行列の要素値である。この目的関数を最小化する際に更新するものは  $A_{\omega}$  及び  $g_{k,t}$  である。しかしこの関数の最小値を解析的に求めることは困難なので、

目的関数の上限を近似的に与え、補助関数法より更新式を導出する。式 (11) において  $\log \sum_k f_{\omega,k} g_{k,t}$  に対して Jensen の不等式を用い、 $\log |A_\omega| = 1/2 \log |A_\omega|^2$  に対して接線不等式を適用すると次式の上限関数を得る。前者の等号成立条件は  $\zeta_{\omega,t,k} = f_{\omega,k} g_{k,t} / \sum_k f_{\omega,k} g_{k,t}$ 、後者の等号成立条件は  $\rho_\omega = |A_\omega|^2$  である。

$$\mathcal{J} \leq \sum_{\omega,t} i_{\omega,t} \left\{ \frac{\sum_k f_{\omega,k} g_{k,t}}{|A_\omega|} + y_{\omega,t} \left( \frac{1}{2\rho_\omega} |A_\omega|^2 - \sum_k \zeta_{\omega,t,k} \log \frac{f_{\omega,k} g_{k,t}}{\zeta_{\omega,t,k}} \right) + C_{\omega,t} \right\} \quad (12)$$

$C_{\omega,t}$  は、アクティベーション行列  $\mathbf{G}$  と全極モデルによるスペクトル重み  $\mathbf{A}$  を更新する際に不必要な成分をまとめて表記したものである。次に上限関数を  $|A_\omega|$  の 2 次形式にするために、 $\tau_\omega$  の周りで  $1/|A_\omega|$  を 2 次のテイラー近似する。等号成立条件は、 $\tau_\omega = |A_\omega|$  である。

$$\mathcal{J} \leq \sum_{\omega,t} i_{\omega,t} \left\{ \sum_k f_{\omega,k} g_{k,t} \left( \frac{1}{\tau_\omega^3} |A_\omega|^2 - 3 \frac{1}{\tau_\omega^2} |A_\omega| + \frac{3}{\tau_\omega} \right) + y_{\omega,t} \left( \frac{1}{2\rho_\omega} |A_\omega|^2 - \sum_k \zeta_{\omega,t,k} \log \frac{f_{\omega,k} g_{k,t}}{\zeta_{\omega,t,k}} \right) + C_{\omega,t} \right\} \quad (13)$$

最後に  $\text{Re}[\theta_\omega^* A_\omega] \leq |A_\omega|$  という不等式を用いることで上限関数が完成する。ここで、 $\text{Re}[\cdot]$  は実部を表し、 $|\theta_\omega| = 1$  である。等号成立条件は  $\theta_\omega = A_\omega / |A_\omega|$  である。

$$\mathcal{J} \leq \sum_{\omega,t} i_{\omega,t} \left\{ \sum_k f_{\omega,k} g_{k,t} \left( \frac{1}{\tau_\omega^3} |A_\omega|^2 - 3 \frac{1}{\tau_\omega^2} \text{Re}[\theta_\omega^* A_\omega] + \frac{3}{\tau_\omega} \right) + y_{\omega,t} \left( \frac{1}{2\rho_\omega} |A_\omega|^2 - \sum_k \zeta_{\omega,t,k} \log \frac{f_{\omega,k} g_{k,t}}{\zeta_{\omega,t,k}} \right) + C_{\omega,t} \right\} \quad (14)$$

以降ではこの上限関数を  $\mathcal{J}^+$  と呼ぶ。

### 3.3.1 アクティベーション行列 $\mathbf{G}$ の更新式

$\mathcal{J}^+$  を  $g_{k,t}$  で偏微分し 0 とおいた式に補助変数の等号条件を代入することでアクティベーション行列の乗法更新式を導出できる。 $\partial \mathcal{J}^+ / \partial g_{k,t} = 0$  より次が成立する。

$$\sum_{\omega} i_{\omega,t} \left\{ f_{\omega,k} \left( \frac{1}{\tau_\omega^3} |A_\omega|^2 - 3 \frac{1}{\tau_\omega^2} \text{Re}[\theta_\omega^* A_\omega] + \frac{3}{\tau_\omega} \right) + y_{\omega,t} (-\zeta_{\omega,t,k} g_{k,t}^{-1}) \right\} = 0 \quad (15)$$

補助変数を代入し整理すると、 $g_{k,t}$  の乗法更新式を得る。

$$g_{k,t} \leftarrow g_{k,t} \frac{\sum_{\omega} i_{\omega,t} y_{\omega,t} f_{\omega,k} / (\sum_{\kappa} f_{\omega,\kappa} g_{\kappa,t})}{\sum_{\omega} i_{\omega,t} f_{\omega,k} / |A_\omega|} \quad (16)$$

### 3.3.2 全極モデルによるスペクトル重みの更新

まず、上限関数  $\mathcal{J}^+$  を全極モデルの係数  $\alpha_q$  で偏微分し 0 とおくことで式 (17) を得る。ここで  $1 \leq q \leq p$

である。

$$\sum_{k=1}^p \alpha_k \sum_{\omega,t} \left[ i_{\omega,t} \left( \sum_k f_{\omega,k} g_{k,t} \frac{1}{\tau_\omega^3} + y_{\omega,t} \frac{1}{2\rho_\omega} \right) \left( \exp(-\pi j \frac{\omega}{\Omega} (k-q)) + \exp(\pi j \frac{\omega}{\Omega} (k-q)) \right) \right. \\ \left. - \sum_{\omega,t} i_{\omega,t} \left[ \left( \sum_k f_{\omega,k} g_{k,t} \frac{1}{\tau_\omega^3} + y_{\omega,t} \frac{1}{2\rho_\omega} \right) \left( \exp(-\pi j \frac{\omega}{\Omega} q) \right) \right. \right. \\ \left. \left. + \exp(\pi j \frac{\omega}{\Omega} q) \right] - \frac{3}{\tau_\omega^2} \sum_k f_{\omega,k} g_{k,t} \text{Re}[\theta_\omega^* \exp(-\pi j \frac{\omega}{\Omega} q)] \right] \\ = 0 \quad (17)$$

式 (17) のうち、以下のように係数を定義する。

$$R_{k,q} = \sum_{\omega,t} \left[ i_{\omega,t} \left( \sum_k f_{\omega,k} g_{k,t} \frac{1}{\tau_\omega^3} + y_{\omega,t} \frac{1}{2\rho_\omega} \right) \left( \exp(-\pi j \frac{\omega}{\Omega} (k-q)) + \exp(\pi j \frac{\omega}{\Omega} (k-q)) \right) \right] \quad (18)$$

$$r_q = \sum_{\omega,t} i_{\omega,t} \left[ \left( \sum_k f_{\omega,k} g_{k,t} \frac{1}{\tau_\omega^3} + y_{\omega,t} \frac{1}{2\rho_\omega} \right) \left( \exp(-\pi j \frac{\omega}{\Omega} q) + \exp(\pi j \frac{\omega}{\Omega} q) \right) \right. \\ \left. - \frac{3}{\tau_\omega^2} \sum_k f_{\omega,k} g_{k,t} \text{Re}[\theta_\omega^* \exp(-2\pi j \frac{\omega}{\Omega} q)] \right] \quad (19)$$

ここで  $R_{k,q}$  を要素に持つ行列を  $\mathbf{R}$ 、 $r_q$  を要素に持つベクトル  $\mathbf{r}$ 、 $\alpha_k$  を要素に持つ  $\boldsymbol{\alpha}$  を用いれば式 (17) は次のようになる。

$$\mathbf{R} \boldsymbol{\alpha} = \mathbf{r} \quad (20)$$

$\mathbf{R}$  はテプリッツ行列であるから、Levinson-Durbin のアルゴリズムを用いて効率的に  $\boldsymbol{\alpha}$  を求めることができる。

各補助変数を代入すると行列  $\mathbf{R}$  とベクトル  $\mathbf{r}$  は次のように書ける。

$$R_{k,q} = \sum_{\omega,t} \left[ i_{\omega,t} \left( \sum_k f_{\omega,k} g_{k,t} \frac{1}{|A_\omega|^3} + y_{\omega,t} \frac{1}{2|A_\omega|^2} \right) \left( \exp(-\pi j \frac{\omega}{\Omega} (k-q)) + \exp(\pi j \frac{\omega}{\Omega} (k-q)) \right) \right] \quad (21)$$

$$r_q = \sum_{\omega,t} i_{\omega,t} \left[ \left( \sum_k f_{\omega,k} g_{k,t} \frac{1}{|A_\omega|^3} + y_{\omega,t} \frac{1}{2|A_\omega|^2} \right) \left( \exp(-\pi j \frac{\omega}{\Omega} q) + \exp(\pi j \frac{\omega}{\Omega} q) \right) \right. \\ \left. - \frac{3}{|A_\omega|^2} \sum_k f_{\omega,k} g_{k,t} \text{Re} \left[ \frac{A_\omega^*}{|A_\omega|} \exp(-\pi j \frac{\omega}{\Omega} q) \right] \right] \quad (22)$$

## 4 評価実験

### 4.1 実験条件

提案手法の有効性を確認するために、音楽信号を対象とした分離評価実験を行った。まず目的音として Microsoft GS Wavetable SW Synth で作成した Oboe

(Ob.), Piano (Pf.), Trombone (Tb.) の3種類を用意した(それぞれの楽譜は[5]を参照されたい)。それらの中から2種類の楽器音を選び、等パワーで混合することにより観測音(分離対象音)を作成した。また教師音として Garritan Professional Orchestra で作成した Ob., Pf., Tb. を使用する。この教師音は、観測音中の各楽器音の音域を含む範囲で、2オクターブにわたり半音階ずつ上昇する24音で構成されている。また教師音および目的音はいずれもサンプリング周波数44.1 kHzであり、FFT点数8192、窓長4096点の矩形窓を512点でシフトさせてスペクトログラムを作成した。また事前学習における学習基底  $F$  の数は100、信号分離における妨害音基底  $H$  の数は50とした。一般化MMSE-STSA推定における振幅圧縮パラメータ  $\beta$  は1.0とした。

客観評価指標として BSS EVAL TOOLBOX[7] による source-to-distortion ratio (SDR) を用いる。SDR は目的音と妨害音の分離度合いと、一連の信号処理によって生じた目的音の線形歪み及び非線形歪みを考慮した推定目的音の品質を表すものであり、値が大きくなるほど精度がよいことを示す。

実験の概要は以下のとおりである。ここでは、混合された2つの楽器音から1つの目的音を分離する問題を考える。分離したい音の教師音は与えられているが、音色が異なっている(なぜならば異なるMIDI音源を使ってそれぞれ生成されているから)。提案手法においては、基底変形の際に用いる全極モデルの次数を1から40までとし、Fig. 1で示した反復処理を8回繰り返した後に得られる変形した教師基底を用いたSNMFによるSDRを計測した。比較対照手法は、2.1節で説明した(基底変形を含まない)SNMFとする。

## 4.2 実験結果

Figure 2は、Pf. と Ob. の楽器音の組み合わせから Pf. の音を分離するという問題に関して、従来のSNMF(SNMF)と提案手法(proposed SNMF with basis deformation)を比較したものである。ここでは、提案手法における全極モデルの次数を1から40まで変化させた時のSDRの値を示している。提案手法が従来のSNMFより良い分離を示す次数が存在することがわかる。ただし、Fig. 2以外の楽器音の組み合わせでは、特定の次数においてSDR値がピークをもつ傾向にあった。

Table 1は従来のSNMFと、提案手法において全極モデルの次数を1から40まで変化させた時のSDRの値が最大になる次数とそのSDRの値を各楽器音の組み合わせに対して記述したものである。これは、組み合わせパターン(例えばOb. & Pf.)のうち、前方に表記されている楽器音(前記例で言えばOb.)を分離する実験の結果を表している。提案手法において、従来のSNMFのSDRを上回る次数が存在することがわかる。ただし、この提案手法における最高性能を示す次数は楽器組み合わせパターンに依存しており、今回の実験ではそれを手動で見つけ出している。本最適次数の自動決定問題に関しては、今後の課題であると言える。

## 5 おわりに

本稿では、SNMFに全極モデルを用いた教師基底の変形成分を反復的に学習する手法を取り入れることにより、従来のSNMFを上回る分離性能が得られる全極

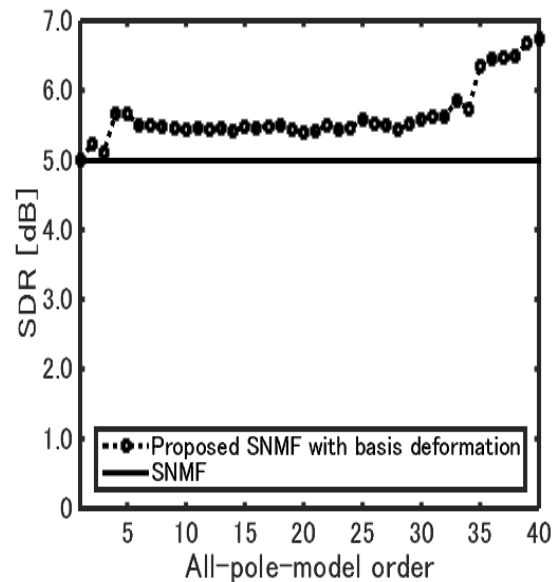


Fig. 2 Pf. と Ob. の混合音から Pf. を分離抽出する場合における SDR

Table 1 各楽器混合パターンにおける分離抽出音の SDR [dB]

	SNMF	Proposed method	Order
Ob. & Pf.	6.278	7.244	26
Ob. & Tb.	2.965	3.136	12
Pf. & Ob.	4.991	6.733	40
Pf. & Tb.	3.888	5.749	40
Tb. & Ob.	1.247	2.375	35
Tb. & Pf.	3.711	4.352	23

モデルの次数が存在することを確認した。

## References

- [1] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Proc. Advances in Neural Information Processing Systems*, vol.13, pp.556–562, 2001.
- [2] 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.
- [3] D. Kitamura, H. Saruwatari, K. Yagi, K. Shikano, Y. Takahashi, K. Kondo, "Music signal separation based on supervised nonnegative matrix factorization with orthogonality and maximum-divergence penalties," *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, vol.E97-A, no.5, pp.1113–1118, 2014.
- [4] D. Kitamura, H. Saruwatari, K. Shikano, K. Kondo, Y. Takahashi, "Music signal separation by supervised nonnegative matrix factorization with basis deformation," *Proc. IEEE 18th International Conference on Digital Signal Processing (DSP2013)*, no.T3P(C)-1, 2013.
- [5] Y. Murota, D. Kitamura, S. Nakai, H. Saruwatari, S. Nakamura, K. Shikano, Y. Takahashi, K. Kondo, "Music signal separation based on bayesian spectral amplitude estimator with automatic target prior adaptation," *Proc. ICASSP2014*, pp.7490–7494, 2014.
- [6] 安良岡直希, 奥乃博, "調波、非調波・音色構造因子分解による音響信号分析と音源分離インターフェースへの応用," *IPSI-MUS*, vol.27, pp.1–8, 2012.
- [7] E. Vincent, R. Gribonval and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Trans. ASLP*, vol.14, no.4, pp.1462–1469, 2006.