

スペクトル包絡保存に基づく NMF による音源分離

浅川智瑛^{†1} 浜中雅俊^{†1†2}

本稿では複数の楽器による混合音源をスペクトル包絡保存に基づく NMF によって個別の楽器の音響信号に分離する方法を提案する。我々の手法は近い音高であればそのスペクトル包絡が周波数方向にシフトするという調波音の一般的な特徴に基づいたものであり、各々の楽器の演奏可能なあらゆる中心周波数において基底を設け、隣接した基底についてその包絡が近似するよう制約を行うことで分離を行う。この操作は各々の楽器の音色を特徴として成分を分離することに相当し、教師なし学習での調波・非調波混合音源の分離を可能とする。性能評価実験として MIDI での混合演奏音源を分離し、SNR を求めた結果リードギターについて約 3.6dB、ドラムについて 6.0dB の分解能を得ることができた。

Separation of music signal sources by NMF based on preservation of spectrum envelope

TOMOAKI ASAKAWA^{†1} MASATOSHI HAMANAKA^{†2}

This paper proposes the method to separate polyphonic music signal into signals of each musical instrument by NMF: Non-negative Matrix Factorization based on preservation of spectrum envelope. Our method bases on a common feature of harmonic signal: spectrum envelopes of musical signal in close pitches played by the harmonic music instrument would be similar. We estimate power spectrums of each instrument by NMF with restriction to synchronize spectrum envelope of bases which are allocated to all possible center frequencies of each instrument. This manipulation means separation of components which refers to tones of each instrument and realizes both of separation without pre-training and separation of signal including harmonic and non-harmonic sound. We had an experiment to decompose mixture sound signal of MIDI instruments into each instrument and evaluated the result by SNR of single MIDI instrument sound signals and separated signals. As a result, SNR of lead guitar and drums approximately marked 3.6 and 6.0 dB and showed significance of our method.

1. はじめに

本稿ではロックなどのバンド形式での演奏における複数の調波楽器・非調波楽器による演奏音が混在した音楽音響信号を各楽器の音響信号へと分離する手法の提案を行う。楽曲データからの特定の楽器の信号の抽出は編集や採譜といった能動的な楽曲の鑑賞方法への応用が期待されており、従来から多くの研究が行われている。音源分離に用いられる技術として近年特に注目されているのが NMF (Non-negative Matrix Factorization: 非負値行列因子分解) である。これは楽曲のスペクトログラムを非負値行列とみなし、複数の基底ベクトル群と時系列アクティベーション群の積として近似する手法であり、音高、リズム、楽器の種類といった一定数の特徴の組み合わせにより成分を説明できる点や、楽器ごとのパワースペクトルの加法性が近似的に成立するといった点で音楽音響信号の性質との親和性がよいため広く研究されている。

これまでの NMF による音源分離には楽譜情報を教師信号として用いて基底の初期値を与える手法[1]や音源分離

とその演奏内容の置換を目的とし、調波・非調波構造と残響情報のモデリングを行う手法[2]、基底の時間ごとの状態変化を定義し、楽器音の非定常性を許容する手法[3]などがあるが、次の2点を同時に達成できる研究はこれまで十分に行われてこなかった。

1) 教師なし学習による分離であること

未知の楽曲について音源分離を行う場合、ユーザが同時に楽譜などの演奏内容に関する情報を与えられていないことが一般的であるため、教師情報を与えずとも分離を達成できなくてはならない。

2) 複数の調波・非調波楽器音が混在した信号を分離すること

NMF による音源分離は解析するデータを基底群の時系列での励起として取り扱うが、それぞれの基底が何を表現しているかについて示唆を与えるものではないため、各楽器の成分としてクラスタリングされた基底とアクティベーションのペアを個別に推定できる枠組みを構築する必要がある。

この2点を解決するため、本研究では楽器ごとに基底のスペクトル包絡が一致するよう制約を行う NMF を提案する。一般的に調波楽器の演奏音は近い音高であれば聴感上同じ音色を持つと仮定すると、時間周波数空間上ではそれぞれの演奏音のスペクトルは同形の包絡を持って周波数

^{†1} 筑波大学大学院 システム情報工学研究科
University of Tsukuba Graduate School of Systems and Information Engineering.

^{†2} JST さきがけ
PRESTO JST

方向にシフトすると考えられる。この仮定に基づき、各調波楽器に発音可能音域での任意の中心周波数に対し割り当てられた基底群が包絡概形を保存するよう制約することで各楽器のパワースペクトログラムの推定を行う。この操作は楽器ごとの音色を特徴として成分を分離することに相当し、1)事前情報なしでの分離と2)調波楽器・非調波楽器の混在信号の分離の双方を実現できる。

2. 関連研究

本節では音源分離において NMF 以前に行われてきた研究および NMF による音源分離において本研究とかわりの深いものについて紹介する。

NMF 以前のアプローチとしては楽器ごとの音響的特徴を利用したモデリングを行い、各種特徴量を EM アルゴリズムなどにより更新していき最適化するパラメトリックな手法が主要であった。後藤らは複数楽器の混合音においてメロディーラインとベースラインを推定する手法として PreFEst を提案した[4]。これはスペクトログラムの各時間フレームを確率密度関数とみなし、EM アルゴリズムによってあらゆる中心周波数についての倍音構造のパワー比を求める手法であり、機械学習による音高推定に多くの影響を与えた。また、糸山らは調波構造・非調波構造モデルを定義し、MIDI 音源による演奏を教師信号として EM アルゴリズムによりモデルを対象信号に近づけていくことでの分離を行った[5]。この研究により提唱されたモデルは多くのパラメータを必要とするものの大きな自由度を持ち、多様なパワースペクトルを表現できたため他の研究でも用いられている。

NMF を用いた研究としては、北野らは調波構造を持つ成分のアクティベーションが時間方向に、非調波成分を持つ成分の基底が周波数方向にそれぞれ滑らかであるという特徴に着目し両者の分離を試みた[6]。この特徴は本研究の目的である調波・非調波混合信号の分離を達成するために有効なものであると考えられるが、調波楽器群を個別に取り扱うためには異なるアプローチの追加を必要とする。また、図1に示すように調波楽器にも非調波成分が存在(ギターのアタック音など)し、非調波楽器にも調波成分が存在(タムの音程感の存在など)する。調波成分と非調波成分を分離するのみではこのような楽器ごとの成分の種類の混在を表現することができないため、統一的なモデル設計を行う必要がある。Raczynski らは音名に対応する数だけ基底を用意し、各々が調波構造を成すよう周波数空間でくし型に初期化することで半音階についての成分の分離を試みた[7]。このモデル設計は調波構造の取り扱いにおいて有効であると考えられるが、やはり調波楽器が複数存在する場合に楽器ごとの成分を分離することはできない。また、半音階のみを取り扱うことによる問題点については後述する。Schmidt ら[8]は楽音の調波構造が対数周波数上

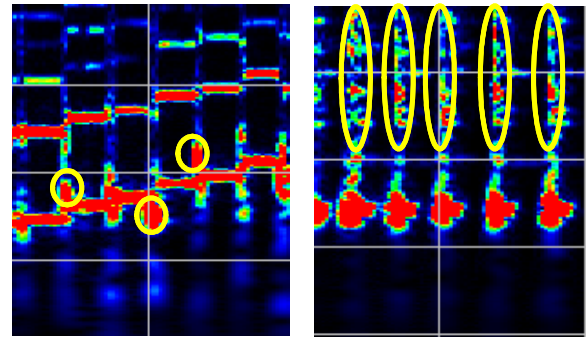


図1：調波楽器に含まれる非調波成分および非調波楽器に含まれる調波成分例

Figure 1: Examples of non-harmonic components of harmonic instrument and harmonic components of non-harmonic instrument

でシフト不変であることを仮定し、中心周波数が増加しても倍音のパワー比が固定されるという仮定のもとに複数の調波楽器の演奏信号の分離を行った。これは調波楽器の音色を特徴量として扱うことに相当するが、1つの楽器の音色があらゆる音高に対し固定されている。発音可能音域が数オクターブある楽器では演奏音の高低により大きく音色が異なるため、この仮定は本研究においては妥当ではないと考えられる。したがって演奏音の高低に従う連続的な音色の変化を許容できることが求められる。

3. 問題提起と解決へのアプローチ

本節ではロックなどのバンド形式での演奏において調波・非調波楽器の混在信号を分離するうえでの課題と解決方法について述べる。一般的なバンド形式での演奏はギター、ベース、ドラム、キーボードなどを用いて行われ、楽器の種類はある程度決まっているが楽曲によって各楽器の音色が大きく異なる。たとえばギターを例にとると、エフェクトをかけないクリーンでの演奏と歪ませたディストーションでの演奏が存在し、イコライザーの使い方でさらに音色が変化する。これは多くの研究で取り扱われるピアノや管楽器、ヴァイオリンなどの混成信号の分離では見られない問題であり、事前に音響的特徴量を与えるのが困難なため教師あり学習を行えない大きな要因となる。この問題を解決するため、本研究では楽器の数と種類を既知とした上で各楽器の音色を学習するよう基底を制約した NMF により音源分離を行う。図2に基底の構造について示す。基底の数についてはあらかじめ楽器によって異なる数を指定する。非調波楽器については調波成分とそれに付随する成分を学習する2種類の基底を設ける。調波楽器についてはその楽器が発音する中心周波数帯に含まれるすべての周波数ビンの数だけ基底を設ける。ここで注意すべきは、半音階に相当するビンだけを与えることが不適當であると

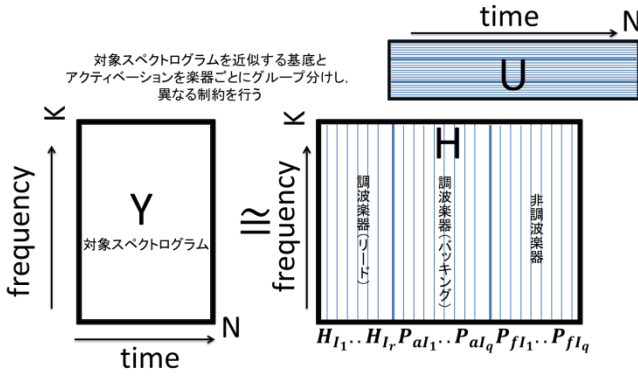


図2 : モデリング概要
Figure 2: Overview of modeling

いう点である。調波楽器はギターにおけるチョーキングなどのように連続的な音高の変化を伴う演奏を行うことがあるため、より高品質な分離を行うためには任意の中心周波数に対応することが妥当だと考えられる。

4. モデルの定式化

本節ではモデルの詳細および定式化について述べる。NMF は推定行列と対象行列の乖離度が単調減少するよう反復計算を行うものであり、乖離度とともに新たに設計したコスト関数を導入し、単調減少させることで目的の制約を行うことができる。ここで観測スペクトログラム Y_{KN} を基底行列 H_{KM} とアクティベーション行列 $U_{m,n}$ の積で近似すると

$$Y_{KN} \cong X_{KN} = H_{KM} * U_{MN} = \sum_m h_{km} u_{mn}$$

$$\sum_k h_{km} = 1 \quad (1)$$

であり、乖離度を補助関数法[9]および KL ダイバージェンスを用いて

$$D(\theta) = \sum_{k,n} (y_{kn} (\log y_{kn} - \sum_m \lambda_{kmn} \log \frac{h_{km} u_{mn}}{\lambda_{kmn}}) - y_{kn} + x_{kn})$$

$$\geq \sum_{k,n} (y_{kn} \log \frac{y_{kn}}{x_{kn}} - y_{kn} + x_{kn}) \quad (2)$$

$$\lambda_{kmn} > 0, \sum_m \lambda_{kmn} = 1$$

とし、減少させるべき目的関数を

$$C(\theta) = D(\theta) + R_p(\theta) + R_h(\theta) + R_b(\theta) + R_l(\theta)$$

と定めると、 $C(\theta)$ を最小化する θ を求める問題となる。以降に各コスト関数およびパラメータ更新式について記述する。

4.1 非調波楽器の分離

非調波構造を主に持つ楽器については1つにつき2種類の基底をペアで割り当てる。打楽器の場合でもタムやスネアなどについては調波的な成分を含むため、一方が非調波成分を、もう一方がそれに付随する成分を学習する。具体的には周波数方向に滑らかな包絡を持つよう制約した基底と、それに同期するようアクティベーションを制約した基底を設ける。非調波楽器 I_{pi} ($i = 1, 2, \dots, q$) に割りあてる基底のインデックスを P_{ahi}, P_{phi} とすると、すべての I_{pi} に対し次のように基底の初期化を行いコスト関数 $R_p(\theta)$ を設ける。

$$h_{km} = 1, (m = P_{ahi}, P_{phi})$$

$$R_p(\theta) = \gamma_{pf} \sum_i \sum_{k=2}^K (h_{kp_{phi}} - h_{k-1_{p_{phi}}})^2 + \gamma_{pa} \sum_i \sum_n (u_{p_{ahi}n} - u_{p_{phi}n})^2 \quad (3)$$

ここで γ_{pf}, γ_{pa} はそれぞれコストの目的関数に対する相対強度を決める重み係数である。

4.2 調波楽器の分離

調波構造を主に持つ楽器についてはまずコードなどの伴奏を演奏するバックキックと主旋律を演奏するリードに種類を分けて考える。これが意味するのはそれぞれ和音と単音を演奏するため、制約条件を変えることで分離を可能にするということである。調波楽器 I_{hi} ($i = 1, 2, \dots, r$) に割り当てる基底のインデックス群を H_{hi} とするとし、すべての I_{hi} に対し次のように基底を初期化したうえで更新を進める。

$$h_{km} = \begin{cases} 1 & (\text{mod}(k - \omega_{lowhi}, \xi) = 0, \xi \in H_{hi}) \\ 0 & (\text{else}) \end{cases}$$

ただし、 $\forall x \in \mathbb{N}, \forall y \in \mathbb{N}$ について

$$\frac{x - \text{mod}(x, y)}{y} \in \mathbb{N}$$

これにより調波楽器 I_{hi} が演奏しうる任意の中心周波数に対し、その n 次倍音をもつくし型の基底を準備する。続いて各調波楽器の音色を学習するためのコスト $R_h(\theta)$ を次のように定める。

$$R_h(\theta) = \gamma_h \sum_{i,j} \sum_{m \in H_{hi}} (h_{\varphi(m,j)m} - h_{\varphi(m-1,j)m-1})^2 \quad (4)$$

なお、 γ_h はコストの目的関数に対する相対強度を決める重み係数、 $\varphi(m, j)$ は $h_{km} \neq 0$ を満たす j 番目のインデックスである。このコストが意味するところは1つの調波楽器の中で隣接した基底の各倍音強度の差が小さくなるよう制約するという点であり、隣接した基底のみを参照することで前節で触れた音色の音高による連続的な変化の学習を可能にした設計になっている。これまでは調波楽器に共通の処理であり、以降にバックキックとリードにおける処理を記述する。

4.2.1 バッキングの分離

バッキングは主に和音を演奏するパートであり、時間フレームごとの全てのアクティベーションに注目したとき複数の中心周波数をもつ可能性がある。これらについて事前に予測を立てることは困難だが、周波数空間で見るとこれらの中心周波数はごく狭い領域に集中すると考えられる。ここで瞬間アクティベーションを確率分布として考えると、理想的な瞬間アクティベーションは周波数軸上での期待値を与える周波数ビン近傍に複数励起することが望ましい。したがって、瞬間アクティベーションの分散を小さくすることによって帰納的な分離を試みる。バッキングに含まれる調波楽器の基底インデックス群を $H_b \subset H_{hi}$ とすると

$$R_b(\theta) = \gamma_b \sum_n \left(\frac{\sum_{m \in H_b} m^2 u_{mn}}{\sum_{m \in H_b} u_{mn}} - \mu^2 \right) \quad (5)$$

$$\mu = \frac{\sum_{m \in H_b} m u_{mn}}{\sum_{m \in H_b} u_{mn}}$$

ここで γ_b はコストの目的関数に対する相対強度を決める重み係数である。

4.2.2 リードの分離

リードは主旋律を演奏するパートであり、特定の中心周波数に相当する基底のみが励起することが望ましい。したがって、前節に倣い瞬間アクティベーションを確率分布として考え、その尖度が増加するよう制約することを考える。リードに含まれる調波楽器の基底インデックス群を $H_l \subset H_{hi}$ とすると

$$R_l(\theta) = -\gamma_l \sum_n \left\{ \left(\sum_{m \in H_l} m^4 u_{mn} - 4\mu \sum_{m \in H_l} m^3 u_{mn} + 6\mu^2 \sum_{m \in H_l} m^2 u_{mn} \right) \left(\sum_{m \in H_l} u_{mn} \right)^{-1} - 3\mu^4 \right\}$$

ここで γ_l はコストの目的関数に対する相対強度を決める重み係数である。

4.3 更新式の導出

前節に記述した目的関数を各コンポーネントで偏微分し、右辺を0とおいて各々について解くことで更新式を得ることができる。

非調波楽器の非調波成分項について

$$\frac{\partial C}{\partial h_{kp_{fhi}}} = 0$$

より、

$$h_{kp_{fhi}} \leftarrow \frac{b_{pf} + \sqrt{b_{pf}^2 + 8\gamma_{pf} h_{kp_{fhi}} \sum_n \frac{y_{kn} u_{p_{fhi}n}}{x_{kn}}}}{4\gamma_{pf}} \quad (6)$$

$$b_{pf} = 2h_{k-1p_{fhi}} \gamma_{pf} - \sum_n u_{p_{fhi}n}$$

非調波楽器の調波成分項について

$$\frac{\partial C}{\partial h_{kp_{ahi}}} = 0$$

より、

$$u_{p_{ahi}n} \leftarrow \frac{b_{pa} + \sqrt{b_{pa}^2 + 8\gamma_{pa} u_{p_{ahi}n} \sum_k \frac{y_{kn} h_{kp_{ahi}}}{x_{kn}}}}{4\gamma_{pa}} \quad (7)$$

$$b_{pa} = 2u_{p_{fhi}n} \gamma_{pa} - \sum_k h_{kp_{ahi}}$$

調波楽器の基底について

$$\frac{\partial C}{\partial h_{km}} \Big|_{m \in H_{hi}} = 0$$

より、

$$h_{km} \leftarrow \frac{b_h + \sqrt{b_h^2 + 8\gamma_h h_{km} \sum_n \frac{y_{kn} u_{mn}}{x_{kn}}}}{4\gamma_h} \quad (8)$$

$$b_h = 2h_{\varphi(m-1,j)m-1} \gamma_h - \sum_n u_{m-1n}$$

調波楽器のバッキング項について

$$\frac{\partial C}{\partial u_{mn}} \Big|_{m \in B} = 0$$

より、

$$u_{mn} \leftarrow \frac{u_{mn} \sum_k \frac{y_{kn} h_{kn}}{x_{kn}}}{\sum_k h_{km} + \gamma_b \frac{\partial \sigma^2}{\partial u_{mn}}} \quad (9)$$

調波楽器のリード項について

$$\frac{\partial C}{\partial u_{mn}} \Big|_{m \in L} = 0$$

より、

$$u_{mn} \leftarrow \frac{-b_l + \sqrt{b_l^2 - 4a u_{mn} \sum_k \frac{y_{kn} h_{km}}{x_{kn}}}}{2a} \quad (10)$$

$$a = -\gamma_l \left\{ 4m^2 \frac{\partial \mu}{\partial u_{mn}} (3\mu - m)A - m^2(m^2 - 4\mu m + 6\mu^2)B \right\}$$

$$b_l = \sum_k h_{km} - \gamma_l m^2(m^2 - 4\mu m + 6\mu^2)A - \frac{12\mu^3}{\sigma^4} \frac{\partial \mu}{\partial u_{mn}}$$

$$+ \frac{6\mu^4}{\sigma^6} \frac{\partial \sigma^2}{\partial u_{mn}}$$

$$A = \frac{1}{\sum_{m \in L} u_{mn} \sigma^4}$$

$$B = \frac{\sigma^2 + 2 \sum_{m \in L} u_{mn} \frac{\partial \sigma^2}{\partial u_{mn}}}{(\sum_{m \in L} u_{mn})^2 \sigma^6}$$

$$\frac{\partial \mu}{\partial u_{mn}} = \frac{\partial}{\partial u_{mn}} \frac{\sum_{m \in L} m u_{mn}}{\sum_{m \in L} u_{mn}} = \frac{m \sum_{m \in L} u_{mn} - \sum_{m \in L} m u_{mn}}{(\sum_{m \in L} u_{mn})^2}$$

$$\frac{\partial \sigma^2}{\partial u_{mn}} = \frac{\partial}{\partial u_{mn}} \left(\frac{\sum_{m \in L} m^2 u_{mn}}{\sum_{m \in L} u_{mn}} - \mu^2 \right)$$

$$= \frac{m^2 \sum_{m \in L} u_{mn} - \sum_{m \in L} m^2 u_{mn}}{(\sum_{m \in L} u_{mn})^2} - 2\mu \frac{\partial \mu}{\partial u_{mn}}$$

なお、基底のスケールによる任意性を防ぐために更新のたびに(1)を満たすよう正規化を行う。

5. 評価実験

本節では提案手法の評価実験とその結果について報告する。性能評価には分離対象となる混合音源と各楽器の個別の音響信号が必要なため、MIDI音源による単一楽器の演奏音源とその混合信号を用いて実験を行った。また、CD音源に対して本手法を用いた結果を示す。

5.1 実験条件

MIDI音源のリードギター (LeadG)、バックギター (BackG)、ベース(Bass)、ドラム (Drums) によって演奏された10秒間の音響信号10曲を分離対象とし、得られた各楽器の信号と個別のMIDI音源信号を用いて次式によるSNRを用いて評価を行う。

$$SNR = 10 \log_{10} \frac{Var(O_t)}{Var(S_t)}$$

O_t : 楽器ごとのMIDI音源信号

S_t : 分離により得られたスペクトログラムと O_t のスペクトログラムの差分を復号して得た音響信号

また、その際の詳細なパラメータを表1に示す。これらは聴感上分離の精度が高くなるよう実験的に適したものを求めたものである。

5.2 実験結果

各楽器のSNRを楽曲ごとに求めた結果と、楽器ごとのSNRの平均を表2に示す。SNRの楽器ごとの平均をみると、特にリードギターとドラムについて高い分離性能を得ることができたことが示されている。これはリードギターにはバックギターとベースに比べ多くの基底を与えているため、より詳細な情報の表現ができたことが1つの要因であると考えられる。またドラムについてはハイハット、スネア、バスドラムを想定して3組の基底を与えているが、これらが明確に分離できているとは言い難いため、これらを個別に推定するためには更なるアプローチが必要であると考えられる。

5.3 CD音源への適用

定量的な評価はできないものの、参考実験として実際の演奏音が収録・マスタリングされているCD音源に対して本手法を適用した結果として、混合音源と分離されたリードギターのスペクトログラムをそれぞれ図3、図4に示す。対象音響信号には先述した実験同様リードギター、バックギター、ベース、ドラムによる演奏が含まれている。MIDIでの実験同様リードギターの多くの成分を分離でき

表1: 評価実験におけるパラメータ

Table 1: Parameters in experiment

Parameters	value
Sampling rate	44100Hz
STFT window	Hamming(8192pts)
overlap	7680
Center frequency range (LeadG)	C4(261.43Hz) -C7(2093Hz)
Center frequency range (BackG)	C3(130.81Hz) -C5(523.25Hz)
Center frequency range (Bass)	C2(65.406Hz) -C4(261.43Hz)
γ_h	16384
γ_{pa}	1
γ_{pf}	131072
γ_l	$1.0 * 10^{-6}$
γ_b	$2.0 * 10^{-3}$

表2: 分離結果[dB]

Table 2: Result of separation[dB]

	LeadG	BackG	Bass	Drums	Ave.
t001	5.442	11.495	-0.662	8.843	6.280
t002	8.973	8.154	-0.595	6.464	5.749
t003	2.921	-2.252	-7.733	-5.217	-3.070
t004	7.908	1.145	1.207	6.695	4.239
t005	2.025	-2.091	4.478	7.075	2.872
t006	6.968	4.097	0.485	6.538	4.522
t007	4.967	-31.206	1.243	10.840	-3.539
t008	-2.956	-0.113	7.732	11.971	4.158
t009	0.913	-2.107	1.602	2.156	0.641
t010	-1.610	-0.537	2.560	4.258	1.168
Ave.	3.555	-1.341	1.032	5.962	

ていることが確認できたが、ピッキングのアタック音の際現に至っておらず、エレキギターの音としての聴感を損なう結果となった。原因としてはエレキギターに与えた基底はすべてくし型に初期化しているため、アタック音の持つ非調波的特性を再現しきれなかったことがあげられる。今後は非調波的特性も含めて近接した音高間で音色が保存されるようなモデルの導入を検討したい。

6. おわりに

本稿ではスペクトル包絡に基づくNMFによるバンド形式での演奏の音源分離手法について述べた。また、本手法の性能を示すために評価実験を行い、その有用性を確認した。本手法の改善の余地として、リードとバックの成分を推定するため瞬間アクティベーションの尖度と分散を

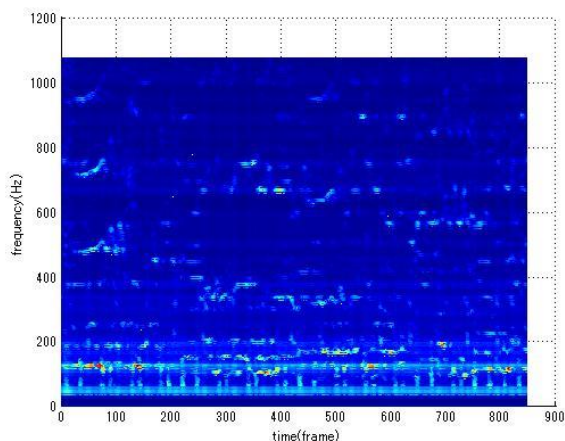


図3：テストデータのスペクトログラム

Figure 3: Spectrogram of test data

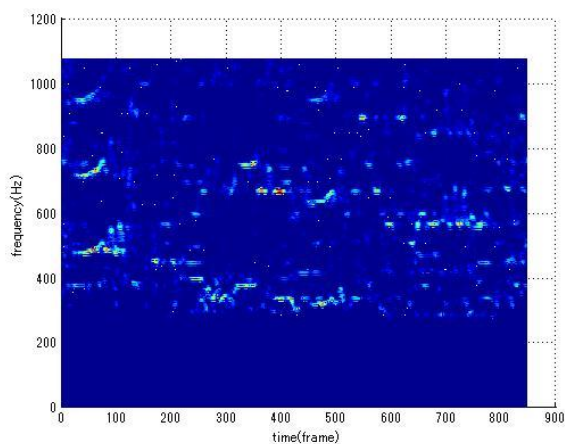


図4：分離したリードギターのスペクトログラム

Figure 4: Spectrogram of separated lead guitar

制約するコスト関数を設けているが、これらが必ずしも一定の値に収束するものではないことがあげられる。それぞれ目的の方向に増加または減少することは確認できているが、計算回数を増やしすぎるとこれらが発散してしまうことが起こりうる。一定の値に収束するよう設計することはすなわち各楽器の信号を分離する指標を定めることにつながるため、事前に指定することが困難であると考えられる。

今後の課題として、前述したコストについて一定値に収束するようなアルゴリズムの設計および音色が類似した楽器による演奏（リードギター2本での演奏など）を分離するための特徴量の考案が求められる。

参考文献

- 1) 糸山克寿, 駒谷 和範, 尾形 哲也, 奥乃 博, “楽譜情報を用いた NMF による音楽音響信号の音源分離”, 情報処理学会全国大会講演論文集 69 巻 2 号 2.163-2.164
- 2) 安良岡 直希, 吉岡 拓也, 糸山克寿, 高橋 徹, 駒谷 和範,

尾形 哲也, 奥乃 博, “フレーズ置換のための調波非調波 GMM・NMF に基づく音源分離・演奏合成”, 情報処理学会論文誌 52(12), 3839-3852, 2011-12-15

3) 中野允裕, 北野佑, ルルージョナトン, 亀岡弘和, 小野順貴, 嵯峨山茂樹, “可変基底 NMF に基づく音楽音響信号の解析”, 情報処理学会研究報告, 2010-MUS-84(10),1-6,2010-02-08

4) Goto, M.: A Real-time Music-scene-description System: Predominant-F0 Estimation for Detecting Melody and Bass Lines in Real-world Audio Signals, Speech Communication (ISCA Journal), Vol.43, No.4, pp.311-329 (2004).

5) 糸山克寿, 後藤真孝, 駒谷和範, 尾形哲也, 奥乃博, “多重奏音楽音響信号の音源分離のための調波・非調波モデルの制約付きパラメータ推定”, 情報処理学会研究報告. 2007(37), 81-88, 2007-05-10

6) 北野佑, 中野允裕, 小野順貴, 嵯峨山茂樹, “スパース性と連続性の制約付き非負値行列分解に基づく 調波音・打楽器音分離”, 日本音響学会春季研究発表会講演集, pp. 833-834, Mar. 2010.

7) Stanisław A. Raczynski, Nobutaka Ono, Shigeki Sagayama, “MULTIPITCH ANALYSIS WITH HARMONIC NONNEGATIVE MATRIX APPROXIMATION”, 情報処理学会研究報告. SLP, 音声言語情報処理 2008(12), 193-198, 2008-02-08

8) Mikkel N. Schmidt, Morten Mørup, “Nonnegative Matrix Factor 2-D Deconvolution for Blind Single Channel Source Separation”, in Proc. ICA2006, pp. 700 - 707, 2006

9) 亀岡弘和, “非負値行列因子分解の音響信号処理への応用,” 日本音響学会誌, Vol. 68, No. 11, pp. 559-565, Nov. 2012.