

楽譜を用いた楽器音分離手法の提案

小林 瑞紀^{a)} 手塚 宏史^{b)} 稲葉 真理^{c)}

概要：技術の発展により，私たちは音楽を聴くだけでなく，音楽を作り編曲することも身近になってきた．多重奏楽曲の中で特定の楽器音のみ抽出し扱うことができれば，それらはより簡単になる．そこで，本研究では多重奏楽曲から特定の楽器音のみを抽出することをめざす．昨今，多重奏楽曲を非負値行列因子分解 (NMF) する楽器音抽出の手法が多く研究されている．今回，楽曲の楽譜情報を用いて NMF の効率を向上させる手法を提案する．

Proposal for the method of Score-based Sound Source Separation

KOABAYASHI MIZUKI^{a)} TEZUKA HIROSHI^{b)} INABA MARI^{c)}

Abstract: As progress of technology, It is becoming familiar for us not only listening to music but making and arranging that. That will become easier if an each instruments sound able to be extracted from polyphonic audio. Then, my research aims at to extract a specific musical instrument sound from polyphonic audio. Today, Many techniques of the musical instrument sound extraction which use Non-negative Matrix Factorization (NMF) of the polyphonic audio are studied. In this paper, the efficient method of NMF using the score information is proposed.

1. はじめに

昨今，技術の進歩により音楽の楽しみ方が，ただ聴くだけでなく自ら作曲をする，編曲を行なうなど多様になってきた．それに伴い，イコライザーを用いて音楽を加工する機会も増えた．音楽で用いられるイコライザーとは，特定の周波数帯域の音量を増減することができる機能，機器である．こと DJ においては，特定の楽器音の音量を増減させるためにイコライザーが用いられている．しかし，実際の楽器音はそれぞれが様々な周波数帯域に及び，また同じ周波数帯域に違う楽器音が重なっている．そこで本研究では，この不便を解消するべく周波数帯域ではなく，楽器のパートごとに音量を増減することができる楽器イコライザーの作成を目的とする．現在，多重音解析による音源分離の分野では非負値行列因子分解 (NMF) が広く用いられている [1]．音響信号のパワースペクトルを加法的と仮定し，非負値行列 Y とみなすと，NMF を用いて周波数 (ス

ベクトル) H と時間 (アクティベーション) U の行列に近似的に分解することができる．すなわち $Y \simeq HU$ と表すことができる．ここで，演奏者が n 人いる楽曲のパワースペクトル Y_{orche} を考える． k 番目の人の演奏のパワースペクトルを $H_k U_k$ で近似できるとき， n 人全員で演奏した元の楽曲のパワースペクトルは， $Y_{orche} = \sum_{k=1}^n H_k U_k$ とみなせる．本研究ではここから個々の $H_k U_k$ を求めたいが，答えが一意に定まらないのは自明である．

そこで，本研究では上記の問題を解決するために初期値として楽譜のデータを用いる事を考えた．曲中出现する楽器ごとの音高の種類とタイミングを楽譜からあらかじめ取得し初期値として用いることで，基底ごとに分解される音高を明確にし，また同じタイミングで違う楽器が鳴った場合にも適切な分解が行われることを可能とする手法を提案する．

なお，本論文の構成は以下のようになっている．セクション 2 では，本研究の根幹技術である非負値行列因子分解 (NMF) などの関連研究を紹介する．セクション 3 では，提案手法である Score-based Sound Source Separation の概要，実装方法を説明する．セクション 4 では，提案手

a) kobayashi.mizuki@ci.i.u-tokyo.ac.jp

b) tezuka.hiroshi@ci.i.u-tokyo.ac.jp

c) mary@ci.i.u-tokyo.ac.jp

法の有効性を実証する実験の手法，結果を示す．セクション5では，提案手法の課題や今後の見通しについて述べる．

2. 関連研究

2.1 非負値行列因子分解 (NMF)

非負値行列因子分解 (Non-negative Matrix Factorization, 以下 NMF) は，要素がすべて非負値の行列 (非負値行列) Y を二つの非負値行列， H と U の積に近似的に分解する多変量解析手法の一つであり，基本概念は Paatero[2] らによって提案された．NMF では，誤差行列 $Y - HU$ の基準として様々な基準が考案され用いられてきた [3], [4], [5]．以下，Lee[6] らによって考案された二乗誤差と I ダイバージェンス基準の反復 NMF アルゴリズムを紹介する．I ダイバージェンスとは一般化 Kullback-Leibler ダイバージェンスとも呼ばれるものである．二乗誤差と I ダイバージェンスはそれぞれ

$$\|A - B\|^2 = \sum_{ij} (A_{ij} - B_{ij})^2 \quad (1)$$

$$D(A||B) = \sum_{ij} (A_{ij} \log \frac{A_{ij}}{B_{ij}} - A_{ij} + B_{ij}) \quad (2)$$

で与えられる．これらより，最小化すべき目的関数は， H と U に依らない項を除くとそれぞれ

$$D_{eu}(Y|HU) = \sum_{\omega,t} (-2y_{\omega,t}x_{\omega,t} + x_{\omega,t}^2) \quad (3)$$

$$D_{kl}(Y|HU) = \sum_{\omega,t} (-y_{\omega,t} \log x_{\omega,t} + x_{\omega,t}) \quad (4)$$

と，表すことができる．ただし，

$$x_{\omega,t} = \sum_k h_{\omega,k} u_{k,t} \quad (5)$$

ここで，Lee らによって考案された補助関数法により，更新式を求めると二乗誤差基準では

$$h_{\omega,k} \leftarrow \frac{h_{\omega,k} \sum_t y_{\omega,t} u_{k,t}}{\sum_t x_{\omega,t} u_{k,t}} \quad (6)$$

$$u_{k,t} \leftarrow \frac{u_{k,t} \sum_{\omega} y_{\omega,t} h_{\omega,k}}{\sum_{\omega} x_{\omega,t} h_{\omega,k}} \quad (7)$$

I ダイバージェンス基準では

$$h_{\omega,k} \leftarrow \frac{h_{\omega,k} \sum_t y_{\omega,t} u_{k,t} / x_{\omega,t}}{\sum_t u_{k,t}} \quad (8)$$

$$u_{k,t} \leftarrow \frac{u_{k,t} \sum_{\omega} y_{\omega,t} h_{\omega,k} / x_{\omega,t}}{\sum_{\omega} h_{\omega,k}} \quad (9)$$

が，得られる．これらの更新式を繰り返すことで，目的関数を減らすべく H と U を更新することができる．

2.2 NMF の多重音解析への応用

P. Smaragdís らは，それまで画像処理の分野で用いられていた NMF を，音楽のスペクトログラムを画像とみなす

ことで多重音解析に適応した [7]．多重音解析に NMF を用いるために，音声信号のパワースペクトルの加法性を仮定して音声信号 Y を基底 (スペクトル行列) H と重み (アクティベーション行列) U の積で近似している．しかし，実際の音声信号のスペクトルは加法的ではない．この問題を解決するために複素 NMF などの手法が提案されている [8]．また，NMF では， Y をより低いランクの行列で近似することを目的としており，観測信号の中で同時に出現する成分についてはそれぞれの成分をひとまとめにした基底が得られやすい．また係数行列の非負性の制約により，係数行列がスパース性を持ち，それぞれ係数が独立な基底が求まりやすい．つまり，楽器音を分離することについて考えると，必ずしも一つ一つの基底ベクトルの組が一つ一つの楽器音に対応しない，という問題を含む．

2.3 楽譜情報を用いた NMF による音源分離

糸山らは，MIDI ファイルに含まれる音符情報を用いて NMF を行い，分離後の信号に含まれるノイズの比率 (signal-noise ratio) を小さくすることに成功した [9]．また，MIDI ファイルを用いることで，NMF のもつ問題点，一つ一つの基底ベクトルの組が一つ一つの楽器音に対応していないことを解決しようとした．事前に MIDI ファイルから音符単位で MIDI メッセージを抽出し，出現するそれぞれの音高ごとに調波構造モデルのテンプレートを作成し，スペクトル行列の初期値として用いた．また，音符のオンセットからオフセットまで定数とした初期値を対応する U の基底に与えた．しかし，楽器ごとではなく音高ごとの調波構造モデルを用いたため，複数の楽器が同じパートを演奏している場合など，分離ができなかった．Sebastian[10] らは，糸山らの手法に加え，楽器ごとに人の聴覚モデルをベースに周波数特性を考え，調波構造モデル [11] テンプレートを作成し，初期値として用いた．また，ビブラートなどの周波数のゆらぎなどにも対応できるモデルを用いた．アクティベーションの初期値には定数を用い，想定する音符の無い時間に相当する要素は 0 で初期化した．また，分離対象である観測信号よりクロマグラフを作成し，隠れマルコフモデルを用いて楽譜と観測信号の時間軸をマッチングさせた．

3. 提案手法

3.1 Score-based NMF の提案

セクション 2.1 より，NMF は観測信号 Y をなるべく低いランクで表現しようとするため，基底と音高や楽器が一对一に対応したような適切な分解をしない．また，与える初期値によって違う結果へと収束する．そこで，楽譜データを用いてこれを解決する．また，楽譜データから音符の情報を得てスペクトルテンプレート，アクティベーションテンプレートを作成するのではなく，事前に楽譜データを

元に音声信号をつくり、それを事前に NMF することでスペクトルテンプレート、アクティベーションテンプレートを得る Score-based NMF を提案した。Score-based NMF では、事前の NMF によって得られるテンプレートが実際の音に近いと、より現実の音に則した分解をすることができる。と考える。

3.2 Score-based NMF の実装

Score-based NMF の実装には、分解対象の楽曲データの楽譜を MIDI ファイルとして用いる。MIDI ファイルは音符がある時間に鳴る、鳴り終わるといった情報だけでなく、楽譜に描かれている、作曲者の想定された演奏をできるだけ再現するような強弱やテンポ、楽器などの情報も含まれている。そのため、それらの情報をできるだけ使うことを考える。まず、あらかじめ MIDI ファイルを読み込んでおき、楽器パートごとに出現する音高を調べる。そして、それぞれの音高ごとにオリジナルと同様のタイミング、強弱、楽器で鳴るよう抜き出した MIDI ファイルを生成する。これを全ての楽器パート、音高ごとに行う。つまり、(楽器パートの数) × (それぞれのパートで出現する音高) だけの MIDI ファイルを生成する。その後、作成した MIDI ファイルをそれぞれソフトウェア音源で鳴らし WAVE ファイルを生成する。そして、各 WAVE ファイルごとにあらかじめ NMF を行い、スペクトルテンプレート、アクティベーションテンプレートを得る。事前にソフトウェア音源を使って得た楽器音を NMF することで、調波構造では表すことができないビブラートなどの楽器音を持つ特徴を初期値のテンプレートとして得ることができる。また、アクティベーションテンプレートにおいてもオンオフを表す二種類の定数だけでなく、リアルなエンベロープ特性をもつテンプレートを得ることができる。ここで得られたスペクトルテンプレートと、アクティベーションテンプレートを H, U の対応する基底の初期値として用いる。

4. Score-based NMF の実験

4.1 概要

提案手法の有効性を確かめるため、モノラル音源の各音高への分離実験を行った。比較対象として、スペクトル行列 H アクティベーション行列 U とともに初期値にランダム値を与えたもの、 H, U とともに初期値として各要素に定数 1 を与えたもの、 H のみに初期値テンプレートを与え U は初期値にランダム値を与えたもの、 U のみに初期値テンプレートを与え H は初期値にランダム値を与えたもの、 H, U とともに初期値としてテンプレートを与え実験を行った。また、第一段階として単音のピアノのみの楽曲を、第二段階として、バイオリンとピアノで演奏されている楽曲を用いた。今回はピアノのみの楽曲としてベートーヴェンの「エリーゼのために」を、バイオリンとピアノの楽曲

としてバッハの「メヌエット第一番」をそれぞれ冒頭 1 フレーズのみ用いた。ここで、第一段階としてピアノを用いたのは、打楽器とは違い明確な音高を持ち、弦楽器のような演奏の仕方による音色のブレが少ないと考えたためである。また、スペクトログラムを得るための短時間高速フーリエ変換の分解窓長は 2048 フレームとし、オーバーライドは 50% とした。窓関数にはハニング窓を用いた。スペクトログラムからの逆フーリエ変換の際、元の音源で得られた位相を与え音声信号の復元を行った。音声信号の復元は各基底ベクトルごとに行った。また、NMF の基底数は出現音高数と同値にした。すなわち、「エリーゼのために」では $K = 9$ に、「メヌエット第一番」では $K = 15$ とした。NMF の距離関数には二乗誤差と I ダイバージェンスを用いた。分解結果として得られた WAVE ファイルの表示には Audacity[12] を用いた。初期値テンプレート作成の際の MIDI ファイルから音高ごとの MIDI ファイルの作成、NMF には MATLAB[13] を用いた。MIDI ファイルからの WAVE ファイルの生成には FluidSynth[14] を用いた。

4.2 スペクトルテンプレートの作成

楽曲に出現する音と同一に近いスペクトルを得るため、同種の楽器のソフトウェア音源で演奏した WAVE ファイルを用いて、NMF することでテンプレートを作成した。テンプレートを作成するための NMF では、基底数を 1 として、反復アルゴリズムの繰り返しも 100 回と制限した。これは、出現する音高の種類は 1 つであるため、1 つの基底で充分網羅することができ、収束も早いと考えたからである。また、得られたスペクトルテンプレート以外の要素は 0 (ゼロ) で初期化した。

4.3 アクティベーションテンプレートの作成

「メヌエット第一番」を用いた実験では、スペクトルテンプレートと同様に、同種の楽器のソフトウェア音源で演奏した WAVE ファイルを用いて、NMF することでテンプレートを作成した。テンプレートを作成するための NMF では、基底数を 1 として、反復アルゴリズムの繰り返しも 100 回と制限した。テンプレート以外の要素を 0 で初期化すると不具合が生じたので、 1^{-10} で初期化した。「エリーゼのために」を用いた実験では、対応する基底の担当する音高が鳴る時間に相当する要素に、 H の初期値にスペクトルテンプレートを、 U の初期値にはランダムな値を用いて事前に「エリーゼのために」を NMF した結果得られた U 内のエンベロープを参考にした値を入力したものをアクティベーションテンプレートとした。

4.4 結果

4.4.1 「エリーゼのために」

単音のピアノ楽曲「エリーゼのために」に対する実験の

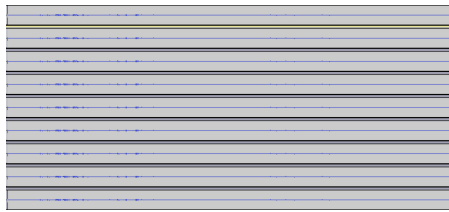


図 1 「エリーゼのために」Iダイバージェンス
 H 初期値: 1 U 初期値: 1
 Fig. 1 "For Elise" / $H : 1 U : 1 / kl$

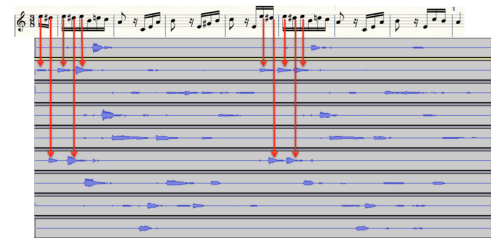


図 2 「エリーゼのために」Iダイバージェンス
 H 初期値:ランダム U 初期値:ランダム
 Fig. 2 "For Elise" / $H : random U : random / kl$

結果は図 1~5 のようになった。今回、二乗誤差より I ダイバージェンス基準のものの方が全体的に良い結果が出たため、I ダイバージェンス基準の結果のみを載せている。これらの図は、 H, U の各基底ベクトル組の積を逆フーリエ変換することで得られた音声信号の波形である。上から、それぞれ 1, 2, 3...9 番目の基底ベクトル組に対応する。例として、図 2 の上部に譜面を載せ、音高が各基底とどのように対応しているか示した。テンプレートは音高の低いものから順に基底の初期値として与えているため、この図では、上から下に行くほど高い音の情報が与えられ NMF をされている。 H か U にテンプレートとして初期値が入ったものは正しくその通りになっている。 H と U に初期値としてすべての要素が 1 の行列を代入した場合以外は、二乗誤差、I ダイバージェンスどちらの距離感数を用いても収束した。図 1 のように初期値として全ての要素が 1 の行列を代入すると、I ダイバージェンスの反復アルゴリズムでは収束させることができないとわかった。 H, U いずれかに初期値としてテンプレートを代入すると概ね音高ごとに分解することができた (図 3~5)。 H, U どちらもランダムな値を初期値として用いると、結果は収束するものの綺麗に音高ごとに分解することはできなかった (図 2)。

また、NMF の繰り返し回数による誤差行列 $Y - HU$ の推移を二乗誤差、I ダイバージェンスで表したものが図 6,7 である。それぞれ、 H, U とともに全ての要素を 1 で初期化したものを青色、 H, U とともにランダムな値を初期値としたものをオレンジ色、 H のみテンプレートを初期値としたものを黄色、 U のみテンプレートを初期値としたものを紫色、 H, U とともにテンプレートを初期値として与えたものを黄緑色で表す。図 8,9 は「エリーゼのために」 H, U 初期値にテンプレートをを用いて音源分離をした結果の第一基底ベクトルから復元した信号のパワースペクトログラム Y_1 を理想としたときの、初期値ごとの第一基底ベクトルの値との誤差行列を二乗誤差、I ダイバージェンスそれぞれで表した図である。

4.4.2 「メヌエット第一番」

バイオリンとピアノで演奏されている楽曲「メヌエット第一番」に対する音源分離の実験の結果は図 10~14 のようになった。今回、二乗誤差より I ダイバージェンス基準の

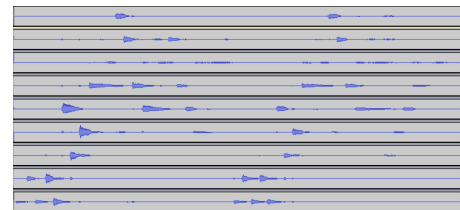


図 3 「エリーゼのために」Iダイバージェンス
 H 初期値:テンプレート U 初期値:ランダム
 Fig. 3 "For Elise" / $H : template U : random / kl$

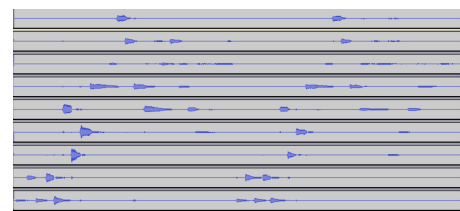


図 4 「エリーゼのために」Iダイバージェンス
 H 初期値:ランダム U 初期値:テンプレート
 Fig. 4 "For Elise" / $H : random U : template / kl$

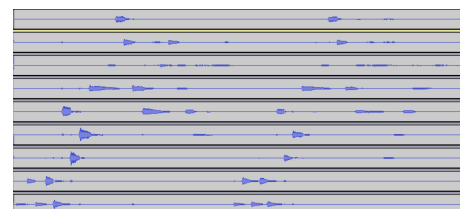


図 5 「エリーゼのために」Iダイバージェンス
 H 初期値:テンプレート U 初期値:テンプレート
 Fig. 5 "For Elise" / $H : template U : template / kl$

ものの方が全体的に良い結果が出たため、I ダイバージェンス基準の結果のみを載せている。これらの図は、 H, U の各基底ベクトル組の積を逆フーリエ変換することで得られた音声信号の波形である。上から、それぞれ 1, 2, 3...15 番目の基底ベクトル組に対応する。テンプレートは音高の低いものから順に基底の初期値として与えているため、この図では、上から下に行くほど高い音の情報が与えられ NMF をされている。また、 H, U とともに初期値としてテンプレートをを用いたものが最も綺麗に音高ごとに分解できた。次いで綺麗に分解できたのは H にランダムな値、 U にテンプレートを初期値として用いたものであった。図 15 はメヌ

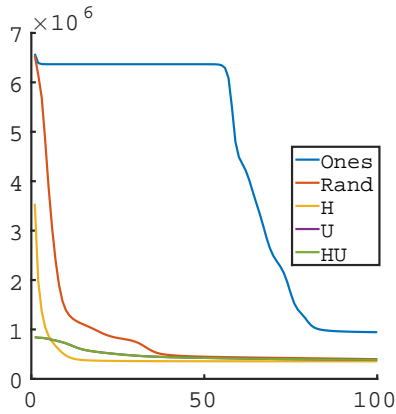


図 6 「エリーゼのために」二乗誤差 D_{eu} 繰り返し回数による推移

Fig. 6 The graph of $\sum D_{eu}$ "For Elise"

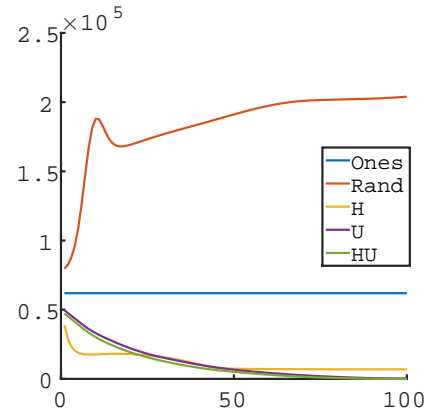


図 9 「エリーゼのために」第一基底ベクトルの理想値との I ダイバージェンス値 D_{kl} 繰り返し回数による推移

Fig. 9 The graph of $D_{kl}(Y_1|H_1U_1)$ "For Elise"

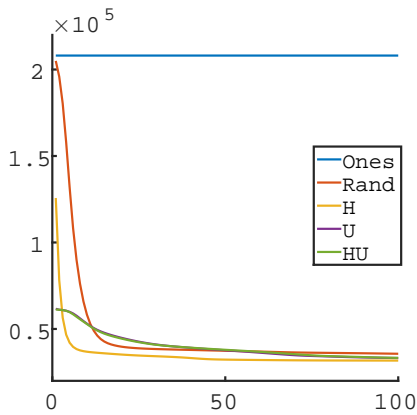


図 7 「エリーゼのために」I ダイバージェンス D_{kl} 繰り返し回数による推移

Fig. 7 The graph of $\sum D_{kl}$ "For Elise"



図 10 「メヌエット第一番」I ダイバージェンス H 初期値: 1 U 初期値: 1

Fig. 10 "Minuet" / H : 1 U : 1 / kl

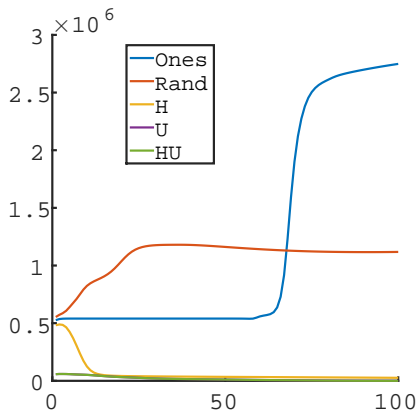


図 8 「エリーゼのために」第一基底ベクトルの理想値との二乗誤差 D_{eu} 繰り返し回数による推移

Fig. 8 The graph of $D_{eu}(Y_1|H_1U_1)$ "For Elise"

エット中に出てくるピアノの「ソ」に相当するテンプレートを初期値として与えた基底の分離結果を拡大したものである。上から、お手本として MIDI をソフトウェア音源で

鳴らした波形, H だけ初期値にテンプレートを与えたもの, U だけ初期値にテンプレートを与えたもの, H, U ともに初期値にテンプレートを与えたものである。一例を拡大したものであるが, これを見ても, H, U ともにテンプレートを与えたもの, U のみテンプレートを与えたもの, H のみテンプレートを与えたものの順に綺麗に分解されていることが分かる。NMF の繰り返し回数による誤差行列 $Y - HU$ の推移を二乗誤差, I ダイバージェンスで表したものが図 16, 17 である。それぞれ, H, U ともに全ての要素を 1 で初期化したものを青色, H, U ともにランダムな値を初期値としたものをオレンジ色, H のみテンプレートを初期値としたものを黄色, U のみテンプレートを初期値としたものを紫色, H, U ともにテンプレートを初期値として与えたものを黄緑色で表す。

4.5 実験結果のまとめ

実験の結果を表 1 にまとめた。ここでの評価は、音高ごとにそれぞれ綺麗に分解されたかどうかである。この表と図 6, 7, 16, 17 を比較、音高ごとにそれぞれ綺麗にバラバラに分解されたかどうかと、誤差行列の値を小さくできたかに

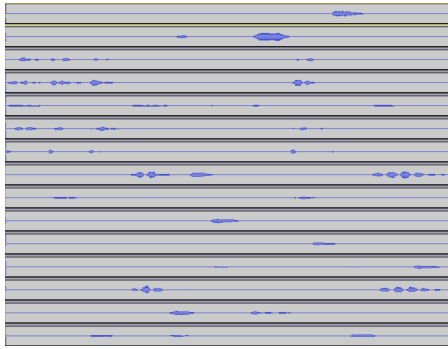


図 11 「メヌエット第一番」I ダイバージェンス
H 初期値: ランダム U 初期値: ランダム

Fig. 11 "Minuet" / H : random U : random / kl

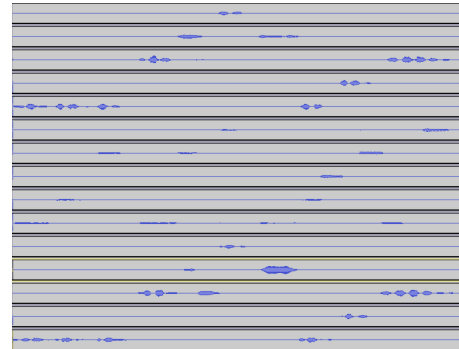


図 14 「メヌエット第一番」I ダイバージェンス
H 初期値: テンプレート U 初期値: テンプレート

Fig. 14 "Minuet" / H : template U : template / kl

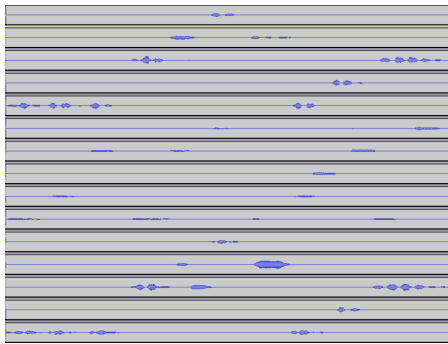


図 12 「メヌエット第一番」I ダイバージェンス
H 初期値: テンプレート U 初期値: ランダム

Fig. 12 "Minuet" / H : template U : random / kl

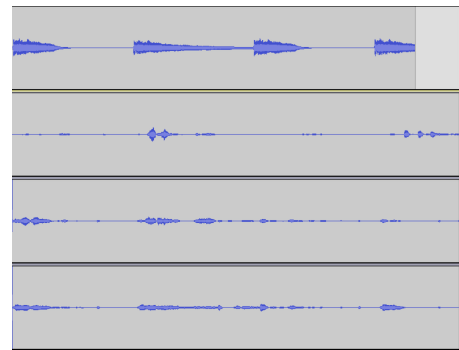


図 15 メヌエット中に出てくるピアノの「ソ」
上から MIDI から生成した wav, H だけテンプレートを与えたもの, U だけテンプレートを与えたもの, HU とともにテンプレートを与えたもの

Fig. 15 Comparing with wav from MIDI, made by using only H template, made by using only U template and made by using H and U template

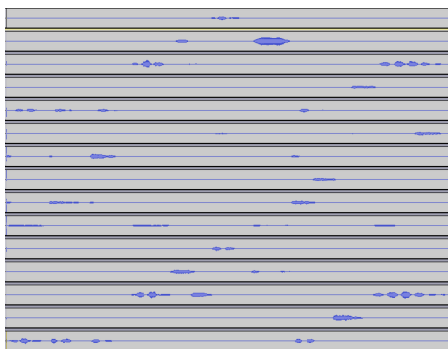


図 13 「メヌエット第一番」I ダイバージェンス
H 初期値: ランダム U 初期値: テンプレート

Fig. 13 "Minuet" / H : random U : template / kl

は相関がないことがわかる。また、波形が綺麗に出ているから楽器音が綺麗に分離できているというわけでもなく、例えば図 13 の二つめの基底は見た目にはノイズも少なく綺麗だが、実際の音はピアノとバイオリンが交じり合った音だった。

全体の基底数を二倍に増やし、ひとつの音高が担当する基底を二つにした、すなわち初期値テンプレートを得る際の NMF で用いる基底の数を 2 に設定して「メヌエット第一番」を分解すると、より綺麗にバラバラに分解され、また誤差行列の値もより小さくなった(図 16,17 の水色)。ま

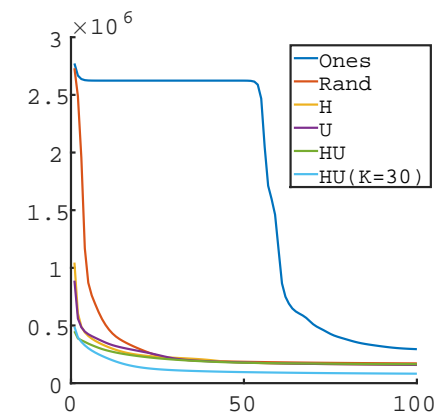


図 16 「メヌエット第一番」二乗誤差 D_{eu}
繰り返し回数による推移

Fig. 16 The graph of $\sum D_{eu}$ "Minuet"

た、一つの基底につき一つの音高を担当していた際には、図 18 の一番上のようにバイオリンのピブラートのような時間ごとに周波数特性が変わるものを再現出来ていなかったが、基底二つに一つの音高を割り当てると、図 18 の上

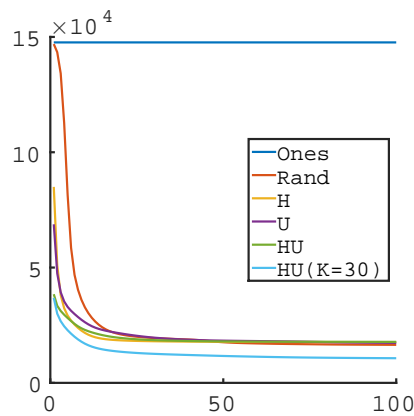


図 17 「メヌエット第一番」Iダイバージェンス D_{kl} の総和繰り返し回数による推移

Fig. 17 The graph of $\sum D_{kl}$ "Minuet"

表 1 「エリーゼのために」, 「メヌエット第一番」を分解する際に与えられた初期値と, その結果

Table 1 A result table

H	U	エリーゼのために	メヌエット第一番
定数 (1)	定数 (1)	×	×
ランダム	ランダム		×
テンプレート	ランダム		
ランダム	テンプレート		
テンプレート	テンプレート		

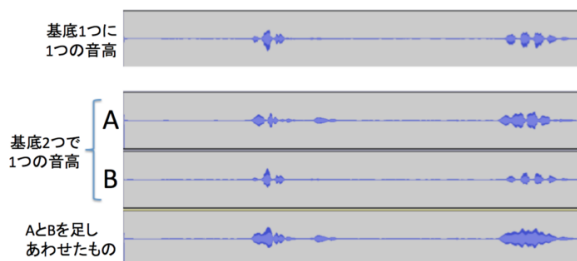


図 18 基底数によるバイオリンの「シ」の音の比較

Fig. 18 Comparing sound of the violin

から 2 番目と 3 番目のように二つの基底で時間ごとに入れ替わることで, ある程度元の音を再現出来た. 図 18 の一番下のものは 2 番目と 3 番目, すなわち同じ音を担当した二つの基底ベクトル組から得た音を足しあわせたものであり, これを見ると一番上の基底が一つのものよりよく再現できているのがわかる. バイオリンなどの弦楽器は, ピブラートによって同じオクターブの同じ音名でも周波数が変異する. そのため, 一つの音高に対し基底を 2 つ以上割り当てるとより再現度が高まる. また, 一つの音高に基底を二つ与えると, ピアノの音はアタック部分のみが別々に別れた. そのため, ピアノの音を分解する際は, アタック部分を担当する基底を音高ごとに用意した方が良く考えられる.

5. まとめ

実際に音が綺麗にそれぞれの基底に分解されていたのは H, U とともにテンプレートを初期値として与えたときだったが, 誤差行列 $Y - HU$ の推移のみに着目すると, 初期値をランダムで与えたものや H のみをテンプレートとして与えたときの方が良い値に収束している. この値は全ての基底 H と重み U を掛けあわせたときの Y からの距離であるため基底ひとつと重みひとつなど, 個々のレベルでの差異は全く加味されていない. そのため, 今後音源分離の研究を進める際には, 分離された個々の信号がどれだけ実利にかなっているかを図る指標が必要である. 例えば, 今回では基底ごとにひとつの音高を割り振ったため, 他の音高成分を含む値をペナルティとして評価する指標などが有用であったと考えられる. 今回の実験結果には入っていないがアクティベーションテンプレートに間違った値を与えると, NMF の過程でスペクトル行列にも悪影響を及ぼし, 単なるランダムな値よりも悪い結果となりうる. そのため今回, アクティベーション行列のテンプレートを作成する過程では手作業が多くあった. そのため, 譜面と楽曲データを高い精度でマッチングさせ, 有用なアクティベーションテンプレートを自動で作成することが今後の課題として考えられる.

参考文献

- [1] 亀岡弘和, 中村友彦, 高宗典玄: 音楽音響信号処理技術の最先端電子情報通信学会誌 Vol.98, No.6 (2015)
- [2] P. Paatero and U. Tapper: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, *Environmetrics*, 5, pp.111-126 (1994)
- [3] I.S. Dhillon and S. Sra: Generalized nonnegative matrix approximations with Bregman divergences, *Proc.NIPS 2005*, pp.283-290 (2005)
- [4] R. Kompass: A generalized divergence measure for nonnegative matrix factorization, *Neural Computation*, 19(3), pp.780-791(2007)
- [5] C. Févotte, N. Bertin, and J.-L. Durrieu: Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis, *Neural Computation*, vol.21, no.3, (2009)
- [6] Daniel D. Lee, H. Sebastian Seung: Algorithms for Non-negative Matrix Factorization, *NIPS*, MIT Press, pp.556-562 (2000)
- [7] Smaragdis, P., Brown, J.C.: Non-negative matrix factorization for polyphonic music transcription Applications of Signal Processing to Audio and Acoustics, *IEEE Workshop* (2003)
- [8] 亀岡弘和, 小野順貴, 柏野邦夫, 嵯峨山茂樹: 複素 NMF: 新しいスパース信号分解表現と基底系学習アルゴリズム, *日本音響学会 2008 年秋季研究発表会講演論文集*, 2-8-13, pp.657-660, (2008)
- [9] 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 第 69 回全国大会*, 2N-1 (2007)

- [10] Ewert, S., Pardo, B., Muller, M., Plumbley, M.D.: Score-Informed Source Separation for Musical Audio Recordings: An overview, Signal Processing Magazine, IEEE (Volume:31 , Issue: 3)
- [11] Stanislaw Andrzej Raczynski, Nobutaka Ono, Shigeki Sagayama: Multipitch analysis with Harmonic Nonnegative Matrix Approximation, Proc. of ISMIR, pp.381-386, (2007)
- [12] Audacity <http://osdn.jp/projects/audacity/>
- [13] MATLAB <http://jp.mathworks.com/products/matlab/>
- [14] FluidSynth <http://www.fluidsynth.org/>