

ドラム演奏の音量バランス習得に向けた 音源分離を用いたリアルタイム叩打音量可視化システムの提案

細谷美月¹ 中村聡史¹ 森勢将雅¹ 吉井和佳²

概要：複数の楽器を同時に演奏するドラムの上達において、主要な3楽器の音量バランスを考慮することが重要であり、基本のビートにおいて、バスドラム、スネアドラム、ハイハットの順に音量が小さくなるように演奏することが理想とされる。この音量バランスの習得において、自主練習をする場合、自分のドラム演奏を客観的に評価することは難しい。また、ドラム演奏を録音することで、自己診断は可能であるが、修正点をふまえて再度演奏・録音し直す方法は効率的とはいえない。これらの問題を解決するため、本研究では半教師あり非負値行列因子分解 (SSNMF) を用いて、ドラム正面に設置した単一指向性のマイクで収録したドラム演奏から、リアルタイムにバスドラム、スネアドラム、ハイハットの各ドラム信号へ分離するとともに、叩打時の音量を推定・可視化することで、ユーザに即時的なフィードバックを与えるシステムを提案する。本研究では、実際のドラム演奏においてシステムを使用してもらった実験を行い、フィードバックからシステムの有用性や改善点が明らかになった。

キーワード：ドラム演奏、練習支援、音量バランス、音源分離

1. はじめに

バンド演奏において、リズムを表現するドラムはバンドの指揮者とも呼ばれている。このドラムは、図1のように複数の楽器によって構成されている。そのうち、使用頻度が高く、演奏の中心となるバスドラム (BD)、スネアドラム (SD)、ハイハット (HI) の3楽器のことをまとめて“3点”と呼ぶ。

このように複数の楽器を同時に演奏するドラムの上達において、リズムの正確さや音色に加えて、3点の音量バランスを考慮することも重要である。この音量バランスは、8ビートのような基本のビートにおいて、バスドラム、スネアドラム、ハイハットの順に音量が小さくなるように (BD>SD>HI) 演奏することが理想とされており、この音量バランスで演奏することによって、まとまりのあるドラム演奏になるとされている。この理想とされる音量バランスが崩れてしまう例として、他の楽器に対してハイハットの音量が大きく、抑揚がない演奏に聴こえてしまう場合や、バスドラムが小さく、安定感のない演奏に聴こえてしまう場合が挙げられる。そのため、ドラム演奏者は理想とされる音量バランスの習得を目指す。

この音量バランスの習得において、ドラム演奏者のうち、ドラムスクールに通うひとや指導してもらった機会があるひとは、演奏を客観的に聴いてもらい、「ハイハットの音量が大きいからもう少し下げよう」といった具体的なアドバイスをもらいながら身につけることができる。しかし、ドラムは管楽器と比べ、見よう見まねで演奏しやすい楽器であり、またピアノやエレクトーンなどと比べ、習いに行く機会も少なく、自己流で練習するひとが多い傾向がある。

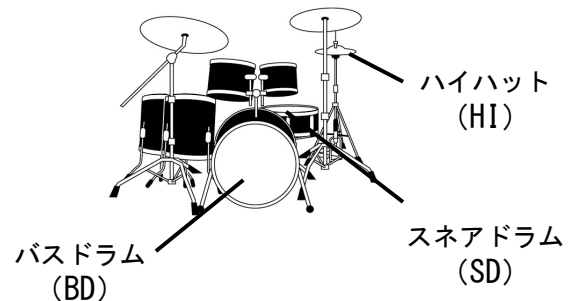


図1 基本的なドラムセット

この場合、客観的に自分の演奏を聴くことができないため、自分のドラム演奏の音量バランスを理解することが難しいという問題が生じる。また、自力で音量バランスの習得を行う場合、演奏を録音し、聴いてみて、修正点をふまえて再度録音するといったステップを繰り返す方法や、各楽器にマイクを設置して音量を測定する方法が考えられるが、こうした方法は時間や手間がかかってしまうという問題がある。一方、MIDI (電子) ドラムで練習することによって、音の強弱を表す数値であるベロシティを確認できるが、叩く時の感覚が生ドラムと大きく異なるため、音量バランスの習得に適しているとはいえない。

我々はこうした、生ドラムでの基本のビート演奏において、ドラム演奏者が自力で理想的な音量バランスを習得する際に生じる、客観的に自分のドラム演奏を聴くことができず、現在の音量バランスがわからないことや、確認に時間や手間がかかるという問題に着目する。また、ドラム演奏者が自力で理想的な音量バランスを容易かつ効率的に習得できる仕組みを実現することを最終的な目標とする。なお、本研究では、ある程度リズムを正確に叩くことができ

1 明治大学
Meiji University
2 京都大学
Kyoto University

るドラム演奏者が、ドラムのみの個人練習で音量バランスの習得を行う場合を対象とする。この理由として、ドラムの上達において、音量バランスを考慮することは、リズムを正確に叩くことができた後のステップであると考えられるためである。

そこで本研究では、ドラムの正面に設置した単一指向性のマイクからドラム演奏を認識し、音源分離によってバスドラム・スネアドラム・ハイハットのみを抽出してそれぞれの音量を計算し、その結果から求めた音量バランスをリアルタイムに可視化してドラム演奏者にフィードバックするシステムを提案する。これにより、今の自分のドラムの音量バランスを一目見て理解することができるため、ドラム演奏者が自力で理想的な音量バランスを容易かつ効率的に習得することを支援できると期待される。

2. 関連研究

2.1 ドラム演奏の練習を支援する研究

ドラム演奏の練習支援にアプローチした研究は様々行われている。

池之上[1]らは微小遅延聴覚フィードバックによって、ドラムスティックの制御を矯正するシステムの提案を行い、システムの使用によって、伸筋を使用する演奏動作に矯正できることが示唆された。触覚提示によって練習支援を行うものとして、Hollandら[2]は、両手両足に装着した装置に振動を提示することによって、複雑なドラムパターンの習得を支援するハプティックプレスレット（ハプティックドラムキット）を提案した。カメラの動作検出を用いた演奏動作の矯正による練習支援も行われており、越智ら[3]はKinectを利用した演奏動作検出によるドラム練習支援システムを提案した。また、本研究と同様にドラムの演奏情報を可視化して提示することにより練習を支援した研究として、安井ら[4]は、MIDIドラム演奏の音の強弱や演奏テンポの変動を演奏時に可視化してユーザに提示する手法を提案している。

しかし、これらの研究で対象とするのは、ドラム演奏のうちリズムや演奏動作に関する支援であり、本研究のような“生楽器ドラムの音量バランス”に着目したものは少ないといえる。

2.2 ドラム演奏の音源分離に関する研究

本研究では、リアルタイムにドラム演奏の音量バランスを算出するために、音源分離を用いる。音源分離とは、複数の音源が混在している音響信号からそれぞれの音源を分離して認識する技術である。これについて、ひとの複数の音を聞き分ける能力を機械で実現することを目的とし、動物の鳴き声やドアの軋む音など特定の音を分離可能としたUniversal Sound Separation[5]や、特定の楽器に関する事前知識を使用せずに音源の調波音と打楽器音を分離する調波

音・打楽器音分離（HPSS：Harmonic/Percussive Sound Separation）[6]のように、目的や条件の違いによって様々な手法が提案されている。

その中でも、本研究で扱うドラム演奏の音源分離については、非負値行列因子分解（NMF：Nonnegative Matrix Factorization）[7,8]や、畳み込み非負値行列因子分解（NMFD：Non-negative Matrix Factor Deconvolution）[9]という手法が提案されている。この手法は、単一チャンネル（モノラル信号）である分離対象の音源の振幅スペクトルを、周波数成分を表現する基底スペクトルと、それぞれの基底スペクトルに対応する時間情報を表現するアクティベーションの行列に分解することによって、音源を分離するアルゴリズムであり、NMFは1フレーム分、NMFDは数フレーム分の時間幅を持った基底行列を学習して行う。また、ディープニューラルネットワーク（DNN：Deep Neural Networks）[10]に基づく手法も提案されており、分離対象の音響信号のスペクトログラムを入力として予測するようなモデルを学習することで、分離を可能としている。こうしたドラム演奏の音源分離の主な活用先は、コンピュータを用いて音響信号から楽譜を自動生成する自動採譜であり、採譜精度を高めるための研究が盛んに行われている[11-15]。

このように、ドラムの音源分離は自動採譜に使用されることが多く、ドラム演奏をしながら音量バランスの習得に用いられるものではない。また、本研究ではシステムを使用する際に演奏するドラムの楽器の基底を事前に登録することが可能であるため、高精度な分離が期待されるNMFを音源分離の手法として採用する。

3. 音量バランス推定手法

3.1 音量バランス推定手法の計算手順

本研究では、ドラムの音量バランスを推定することを目指す。そこで、単一指向性のマイクにより入力されたドラム演奏から、それぞれの楽器の音に分離し、各々の音量を推定する。その際に、音響信号における振幅から算出した、ひとが音圧として感じる指標となるRMS値を音量の計算に使用する。以下にその計算手順を示す。

- (1) 単一指向性のマイクからドラム演奏を認識（図2）
- (2) 直近の約1小節分（4/4拍子）の秒数ごとに、ドラム演奏をハイハット・スネアドラム・バスドラムのみを音源に分離（図3）
- (3) 分離した音源それぞれについてオンセット（音の出だしのタイミング）を検知[16]（図4）
- (4) 検出されたオンセットの前後 t_1 秒分のフレームからピーク（最大振幅）を求め、閾値 T を超えたピークを各楽器の正しい打叩タイミングとして採用（図5）

- (5) 各ピークを境にした t_2 秒分のフレームから RMS 値を算出し、平均値を算出
- (6) (5)で算出された RMS の平均値について、ひとがドラム演奏を聴いた際に感じる音量バランスに近づくように、補正值 (ω_{HI} , ω_{SD} , ω_{BD} とする) をかけあわせ、各楽器の音量とする
- (7) (6)で求めた各楽器の音量を、割合で示したものを音量バランスとする

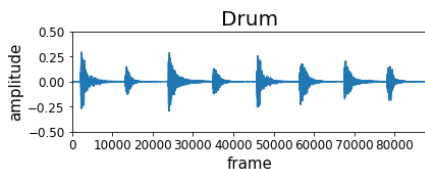


図2 マイク入力から認識したドラム演奏の波形

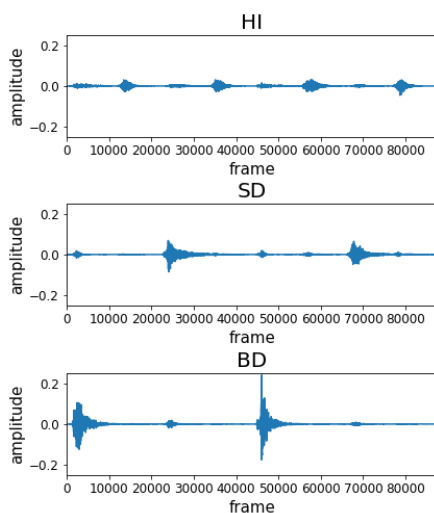


図3 音源分離後の波形

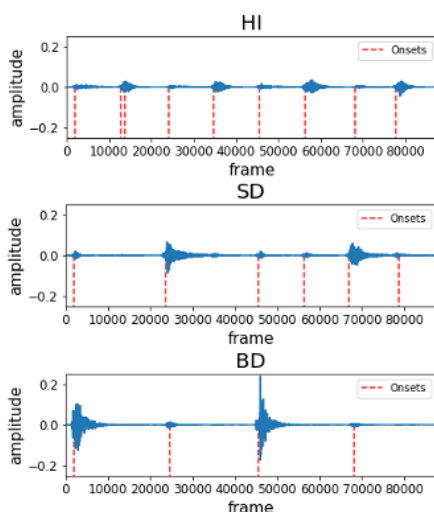


図4 オンセット検出を行った波形

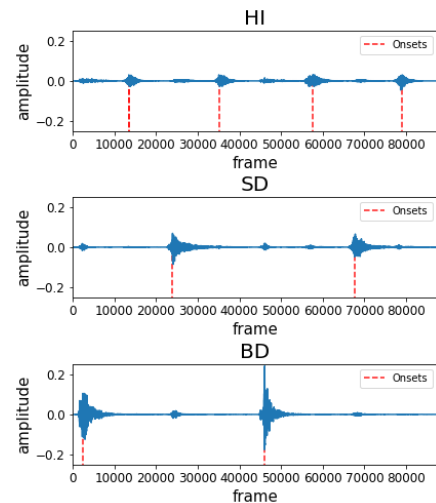


図5 ピーク検出を行い、閾値を超えたものの波形

ここで、(2)について、直近の約1小節分(4/4拍子)の秒数のフレームごとに音量バランスを計算する理由としては、ひとがドラム演奏を聴いて、その音量バランスを判断するにおいて必要最低限の長さであると考えたためである。また、(5)について、RMS値を各楽器の音量として採用した理由としては、ピークは瞬間的な音の大きさを示す数値であるのに対し、音の持つエネルギーを平均して算出される値であるRMS値は、ひとの音圧知覚を考慮した指標であると考えたためである。なお、今回RMS値についてはdB単位に変換せず、そのままの値を使用した。また、(6)について、RMS値に補正值をかけあわせた理由としては、各楽器で検出された発音タイミングのRMS値の平均から算出した音量バランスと、ひとがドラム演奏を聴いた際に感じる音量バランスには差があり、チューニングを行う必要があるためである。

3.2 音量バランス推定に用いた音源分離について

音源分離の手法は様々あるが、今回は、NMFによって分離したい音源の基底スペクトル(周波数成分を表す行列)を求め、その基底を用いて、分離したい音源の基底を固定化することができるSSNMF[17]によって、各楽器の音源へ分離する手法を採用した。これらは、ある非負値行列を2つの非負値行列に分解するアルゴリズムであり、音声データから得られる振幅スペクトログラムを非負値行列とみなし、周波数成分を表す基底行列と、時間情報を表すアクティベーション行列の積で表すことができるという性質を利用した手法である。また、SSNMFを採用したのは、音色や響き具合のような楽器のもつ固有差に対応できるようにするためである。あらかじめ基底を固定化することができるSSNMFで音源分離を行うことで、システムで使用するドラムの楽器の音を録音したものから作成した基底を使用することができ、楽器の固有差に対応できると考えた。

次に、本研究で採用した音源分離の手法について詳しく手順を追って説明する。まず、NMFで使用する基底の作

成を行うため、ハイハット・スネアドラム・バスドラムの単音を、クリッピング（録音時の入力大きい場合に、信号波形のピーク部分が一定のレベルで飽和し潰れてしまうこと）しないように録音し、切り出し、正規化を行いプログラムに登録する。その後プログラム内で、各楽器の録音データについて、短時間フーリエ変換し、得られた振幅スペクトログラムに NMF をかけることで、基底行列を得る。そして、マイクから認識したドラム演奏の録音データをリアルタイムに短時間フーリエ変換し、得られた振幅スペクトログラムに、事前に得られた各楽器の基底を適用した SSNMF をかける。これにより、ハイハット・スネアドラム・バスドラムのみの基底行列と、アクティベーション行列を得ることができ、これらをかけあわせた値を逆フーリエ変換することによって、楽器ごとに分離した音源を得ることができる。

4. 提案システム

本研究では、基本のビートのドラム演奏において、ドラム演奏者が自力で理想的な音量バランスを習得する際に生じる、現在の音量バランスがわからないことや、確認に時間や手間がかかるという問題に着目し、ドラム演奏者が自力で理想的な音量バランスを容易かつ効率的に習得可能とすることを目的とする。ここで、各楽器にマイクを設置して音量を確認する方法も考えられるが、様々な機器や準備が必要であり、容易とはいえない。

そこで、ドラムの正面に設置した単一指向性のマイクからドラム演奏を認識し、音源分離によってバスドラム・スネアドラム・ハイハットのみを抽出した音源からそれぞれの音量を計算し、その結果から求めた音量バランスをリアルタイムに可視化するシステムを提案する。このシステムの使用想定環境を図 6 に示す。ドラムの正面から離れた位置にマイクを設置し、PC をドラム演奏者が演奏しながら画面を確認できる位置に設置したうえで、マイクと PC をオーディオインタフェースで接続する。また、PC 上でシステムを動作させ、音量のバランスを計算するとともに結果を可視化する。

4.1 実装

システムは、音響処理については Python、可視化については Processing を用いて実装し、PC 上で操作を行うものである。

ユーザに提示するシステム画面のイメージ図を図 7 に示す。音量バランス可視化ゾーンにおける赤、緑、青の棒グラフは、リアルタイムに認識したドラム演奏の音声データから計算したハイハット、スネアドラム、バスドラムの音量バランスを表す。また、図右側の目標音量バランス設定バーの操作によって各楽器の割合を設定する。さらに、この音量バランスの割合の数値と連動し、音量バランス可視

化ゾーンにおける目標音量バランス目安バーの値も変化する。なお、画面右下のテンポ（BPM）設定ノブにより、練習するテンポを調整可能とする。

3 章で述べた音量バランスの推定において、音源分離した各楽器の音源に対してピークを求める際に、各楽器の正しい打叩タイミングを認識できるようにするため、 $t_1 = 0.15$ 、 $T = 0.8$ とした。また、騒音計での騒音の測定における Fast 特性[18]の時定数が 125ms であることから、 $t_2 = 0.125$ とした。

次に、ひとがドラム演奏を聴いた際に感じる音量バランスに近づくように RMS 値の平均をチューニングするため、今回は適切な ω_{HI} 、 ω_{SD} 、 ω_{BD} を主観により設定した。具体的には、オーディオインタフェース経由で録音する特性上、ハイハットの入力波形がかなり小さく、バスドラムはやや大きく算出されていたため、 $\omega_{HI} = 5$ 、 $\omega_{SD} = 1$ 、 $\omega_{BD} = 0.8$ （ハイハットについて 5 倍、バスドラムについて 0.8 倍）とした。

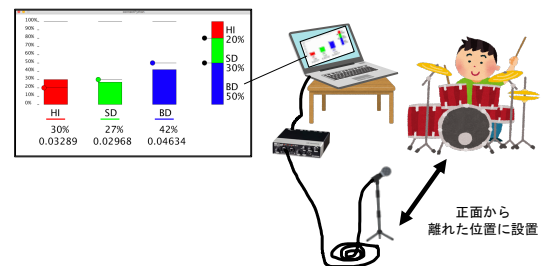


図 6 システム使用想定環境

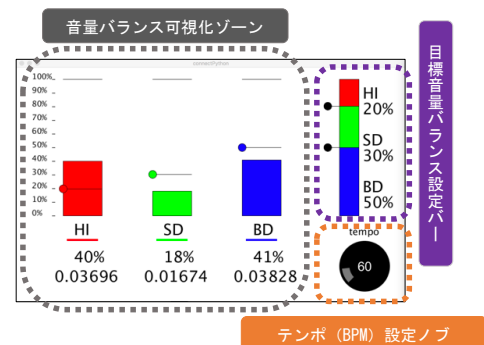


図 7 システムのイメージ図

4.2 利用方法

システムの利用手順において、まずユーザ（ドラム演奏者）は、図 6 のようにドラムの正面から離れた位置に単一指向性のマイクを設置する。次に、PC からシステムを起動し、習得したいドラム 3 点における音量バランスと、演奏するテンポを設定する。その後、ユーザはドラム演奏を開始し、4/4 拍子で 1 小節分（設定したテンポによって変化、BPM60 の場合 4 秒間）ごとに更新される音量バランスが可視化されたシステム画面（図 7）を見ながら演奏を変化させていく。なお図 7 の場合、ハイハットの音量を示

す赤いバーが、設定バーを上回っており、最初に設定した音量バランスに対してハイハットの音量が大きいことがわかる。そのため、ユーザがドラム演奏のハイハットの音量を小さく叩くように演奏を修正することによって、設定した目標に近い音量バランスで演奏することが可能となる。

5. システム使用実験

提案システムによって、ドラム演奏者が自力で理想的な音量バランスを容易かつ効率的に習得することを支援できるかどうか調査するとともに、システムの改善点やフィードバックを得るため、実際にシステムを使用しながら生ドラムを演奏してもらう実験を行った。本実験の実験協力者はドラム歴が5年以上の20代の男性4名である。実験システムを利用している様子を図8に示す。

実験では、事前準備として、システムと機器(PC, マイク, オーディオインタフェース)のセッティングと音源分離に使用するドラムの各楽器(ハイハット・スネアドラム・バスドラム)の基底の作成を実験監督者である筆頭筆者が行った。基底の作成について、今回は実験監督者の指示のもと、実験協力者に各楽器について1音ずつ通常の音量(実験協力者の判断に任せた)で叩いてもらい、録音を行った。その後、録音した音源について、オンセットで切り出し、正規化を行った各楽器の音源にNMFをかけ、基底を作成し、プログラムに登録するところまで実験監督者が行った。

準備終了後、実験協力者にシステムの説明を行い、その後実験に関する指示を行った。指示の内容としては、音量バランスの目標を設定し、その音量バランスになるように演奏をすること、演奏は裏打ち8ビートのみ行うことを併せて伝えた。この裏打ち8ビートとは、8分音符をベースとしたリズムで、各楽器の音が被ることなく演奏するものであり、ドラム演奏において基本となるビートパターンの1つである(図9)。この理由として、今回はシステムの精度ではなく有用性の調査に焦点を当てているため、音源分離しやすいと判断した裏打ち8ビートを採用した。

その後、実験協力者にシステムの使用を開始してもらった。具体的には、システムで実験協力者が習得したい目標の音量バランスを設定してもらい、提示される音量バランスの可視化を見ながら演奏を目標の音量バランスに近づけていくタスクを行ってもらった。システム使用中は録音・録画を行っており、実験協力者が可視化を見て目標の音量バランスで叩いていると判断した時点でシステムの使用を一度終了し、その後録音を聴いてもらうことで、目標の音量バランスで演奏できたかどうかを判断してもらった。

タスク終了後に、目標の音量バランスで演奏できたかどうか、システムの使いやすさなどについてのアンケート、インタビューを行った。具体的な内容を表1に示す。



図8 実験の様子

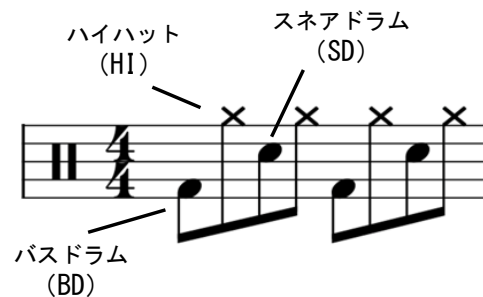


図9 裏打ち8ビートの楽譜

表1 実験で使用したアンケート項目一覧

	質問内容	回答方式
Q1	システムの満足度	5段階評価 (-2~2)
Q2	今後システムを使用したいか	5段階評価 (-2~2)
Q3	(録音を聴いて)設定した音量バランスで演奏できたか	5段階評価 (-2~2)
Q4	システムを使用した感想	自由回答

6. 結果と考察

実験で行ったアンケートの結果をまとめたものを以下に示す。表2に、Q1「システムの満足度」、Q2「今後システムを使用したいか」、Q3「(録音を聴いて)設定した音量バランスで演奏することができたか」について、5段階(-2~2)のリッカート尺度で評価してもらった値を示す。この表より、評価値の平均について、Q1「システムの満足度」は0.75、Q2「今後システムを使用したいか」は1、Q3「(録音を聴いて)設定した音量バランスで演奏することができたか」は1.25という結果であった。

また、表3、4に、Q4「システムを使用した感想」への回答から得られたフィードバックを示す。表3については、システムに対する肯定的な意見、表4については、改善点に関する意見をまとめた。

表2より、評価値は全て正の値となり、また表3よりシステムに対する肯定的な意見が得られたため、リアルタイムにドラム演奏の音量バランスを提示するという本研究の

提案手法によって、目標とする音量バランスを習得できる可能性が示唆された。

しかし、表4に示したように、システムの改善点に関する意見が様々得られたため、これらの意見をふまえて、今後改善していく必要がある。

まず、音量バランスの可視化について、約1小節分の演奏についての計算が終了したタイミングで結果を切り替えていたが、演奏のタイミングと連動しておらず、「いつ演奏した分の音量バランスが反映されているのかが分かりづらい」という意見がみられた。そのため今後は、メトロノーム機能をシステムに組み込み、そのタイミングと連動させて可視化を切り替えることで、タイミングの同期を図る。また、「数秒前の演奏から計算された音量バランスのフィードバックだとわかりづらい」という意見もみられたため、今回用いたような割合を棒グラフで示す以外の手法についても検討していく。具体的に、可視化を切り替える間隔を短くした手法や折れ線グラフで各楽器の音量の推移を可視化する手法などと比較し、適切な手法を見つけていきたい。

また、各楽器の音量の細かい変化に対応できていないことがわかった。これは、音源分離によって算出された各楽器のRMS値の平均を元に計算した音量バランスと、実際にドラム演奏をドラムから少し離れた位置から主観的に聴いた時の感覚的な音量バランスが対応するようなチューニングを、筆頭著者の主観で行ってしまっていたことが原因と考えられる。そこで今後は、ひとにドラム演奏を聴いて

もらい、音量バランスを判断してもらう実験を行い、音源分離によって算出された各楽器のRMS値の結果と対応づけることによって、適切なチューニング方法を見つけていく予定である。

本システムは生ドラムでの個人練習における基本のビート演奏を対象としているが、今回は分離の都合上、図9のように各楽器の音が被らない裏打ち8ビートのみを演奏対象としていた。ほとんどのドラム演奏のビートは、各楽器の被るタイミングが存在するが、このような場合、音源分離にマスキング効果（妨害音によってある音の最小可聴値が上昇する現象[19]）の影響が出てしまう。そこで、このような二つ以上の音が重なったとき、片方がかき消されて鳴っているのに聴こえないまたは聴こえにくくなるという現象が音源分離に与える影響について調査し、裏打ち8ビート以外の演奏も扱えるように検討していく必要がある。また、「他の楽器と一緒に演奏している時の音量バランスも知りたい」という意見もみられたため、最終的にはバンド練習での使用にも対応させたいと考える。

その他の改善点として、音源分離で使用する基底の取得について、今回は筆頭著者がシステム使用前に、録音、正規化や切り出しの編集、プログラムへの登録を行ったが、これについてもシステムに組み込むことで、より効率的なシステムにできると考える。また、UIについてもフィードバックを参考に改善していく予定である。

表2 Q1,Q2,Q3 への回答 (-2~2 の5段階評価)

実験協力者	Q1 システムの 満足度	Q2 今後システムを 使用したいか	Q3 設定した音量バランスで 演奏できたか
A	1	2	1
B	1	0	2
C	0	1	1
D	1	1	1
平均	0.75	1	1.25

表3 Q4 「システムを使用した感想」への回答よりシステムに対する肯定的な意見

分類	回答内容
精度	演奏（の音量バランスが）反映されていてすごいと思った
	わざと大きくしたり小さくしたり叩いたのが反映されていた
可視化方法	意図的に音量バランスを変えた後、すぐ（可視化に）反映された
感想	自分の演奏について客観的に分析することができ、参考になった
	自分の音量バランスを、具体的な数値やグラフで確認することができた
	客観的な数字で自分の音量バランスを見て良かった
	相対的に、この楽器が小さいのだなというのが分かって良かった
	テンポに気をつけながら音量バランスを意識する練習になった （リズムを）正確にする以外にも、割合を意識して叩くことは重要だと思った

表4 Q4 「システムを使用した感想」への回答よりシステムに対する改善点に関する意見

分類	回答内容
精度	スネアを強く叩いても、反映されていない感じがした
	バスドラムが（実際の大きさより）大きく表示されている気がした
	自分の演奏と表示が大まかには連動していたが、細かく変化させた時に連動していないように感じるがあった
	叩いているのに0%になってしまうことがあった
	SDとBDが混同して結果がでてしまっていると感じたことがあった
	音色に左右されてしまう気がした（リムの有無・叩く角度）
	テンポをあげる（早くする）と精度が落ちる気がした
可視化方法	バランスでなく、個々の楽器の音量が見れると良いと感じた
	音量バランス（の可視化）が、いつ叩いた分の結果か分かりづらかったため、切り替わるタイミングを1小節の頭ぴったりにしてほしい
	数秒前の演奏の結果がフィードバックされると、今演奏しているのと違う場合があるので、もう少しリアルタイム感が欲しい
UI	（音量バランス設定について）100%だとどちらかを上げるとどちらかを下げることになるため、比率を自由に入力できるようにしてもいいのかなと思った
	（BPM変更ノブについて）マウスだと値がスムーズに設定できなかつたので、PCの矢印キーで設定できたらいいなと思った
難しさ	メトロノームのテンポを意識しながら、可視化を見て演奏の音量を調整するのが難しかった
使いやすさ	少しシステムを使って慣れないと、どのぐらい演奏を変えたら音量バランスの表示に反映されるかわからなかった
感想	普段邦ロックを叩くことが多く、オープンハイハットや早い曲調が多いため、そこが対応できれば、より使いたいと思う
	バンド全体（他の楽器と合わせている場合）での音量バランスも知りたい

7. おわりに

複数の楽器を同時に演奏するドラムの上達において、基本となる楽器であるバスドラム、スネアドラム、ハイハットの音量バランスを考慮することは重要であり、基本のビートにおいて、バスドラム、スネアドラム、ハイハットの順に音量が小さくなるように演奏することが理想とされている。しかし、生ドラムでの基本のビート演奏において、ドラム演奏者が音量バランスを習得する際に、ドラム演奏者は客観的に自分のドラム演奏を聴くことが出来ないため、現在の音量バランスがわからないという問題がある。また、演奏を録音し、聴いてみて、修正点をふまえて再度録音するなどの方法が考えられるが、確認に時間や手間がかかるという問題がある。本研究ではこうした問題に着目し、基本のビートのドラム演奏において、ドラム演奏者が自力で理想的な音量バランスを容易かつ効率的に習得することの支援を目的とし、ドラムの正面に設置した単一指向性のマイクから認識したドラム演奏を、バスドラム・スネアドラム・ハイハットに音源分離したうえで、それぞれの音量を計算し、音量バランスをリアルタイムに可視化して提示す

るシステムを提案・実装した。また、実際にドラム経験者にシステムを使用してもらい、フィードバックを得る実験を行った。

結果として、提案手法により、容易かつ効率的に目標とする音量バランスを習得できる可能性が示唆されたが、演奏のタイミングと連動させて音量バランスの可視化を切り替えることや、音量バランスの計算精度をより高めることなど、改善点も多くみられた。そこで今後は、得られたフィードバックを元に、システムをよりよいものにしていく予定である。

また今後は、今回筆頭筆者が主観的に設けた基準によって行った音量バランスの計算について、ドラム経験者に演奏を聴いてもらい音量バランスを評価してもらう実験を行うことで、音源分離によって算出された音量バランスと、実際にひとがドラムの正面から演奏を聴いたときに判断する音量バランスの間のチューニング方法について考えていく予定である。さらに音量バランスの可視化方法についても、今回用いた棒グラフ以外の手法と比較する実験を行い、適切なものを見つけていきたいと考えている。

謝辞 本研究の一部は、JST ACCEL(グラント番号 JPMJAC1602)の支援を受けたものである。

参考文献

- [1] 池之上あかり, 小倉加奈代, 西本一志. 微少遅延聴覚フィードバックを用いたドラム演奏練習支援システムの基礎的検証. 情報処理学会 インタラクション 2012 論文集 (情報処理学会シンポジウムシリーズ), 2012, vol. 2012, no. 3, pp. 533-538.
- [2] Holland, S., Bouwer, A. J., Dalgelish, M., Hurtig, T. M.. Feeling the beat where it counts: fostering multi-limb rhythm skills with the haptic drum kit. Proceedings of the 4th International Conference on Tangible and Embedded Interaction, 2010, pp. 21-28.
- [3] 越智洋司, 平野光正, 井口信和. Kinect を利用した演奏動作検出によるドラム練習支援システムの提案. 教育システム情報学会誌, 2017, vol. 34, no. 1, pp. 32-43.
- [4] 安井希子, 三浦雅展. ドラム基礎演奏の練習支援システム(システム論文特集号). 日本音響学会誌, 2015, vol. 71, no. 11, pp. 601-604.
- [5] Kavalerov, I., Wisdom, S., Erdogan, H., Patton, B., Wilson, K., Le Roux, J., Hershey, J.. Universal sound separation. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2019, pp. 175-179.
- [6] Ono, N., Miyamoto, K., Kameoka, H., Le Roux, J., Uchiyama, Y., Tsunoo, E., Nishimoto, T., and Sagayama, S.. Harmonic and Percussive Sound Separation and its Application to MIR-related Tasks. Advances in Music Information Retrieval, 2010, vol. 274, pp. 213-236.
- [7] Lee, D. D. and Seung, H. S.. Learning the parts of objects with nonnegative matrix factorization. Nature, 1999, vol. 401, pp. 788-791.
- [8] Cichocki, A., Zdunek, R., Phan, A.H. and Amari, S.. Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation. Wiley, 2009.
- [9] Smaragdis, P.. Non-negative Matrix Factor Deconvolution; Extraction of Multiple Sound Sources from Monophonic Inputs. International Conference on Independent Component Analysis and Signal Separation (ICA), 2004, pp. 494-499.
- [10] Grais, E. M., Sen, M. U., and Erdogan, H.. Deep neural networks for single channel source separation. ICASSP, 2014, pp. 3734-3738.
- [11] Ueda, S., Shibata, K., Wada, Y., Nishikimi, R., Nakamura, E. and Yoshii, K.. Bayesian drum transcription based on nonnegative matrix factor decomposition with a deep score prior. ICASSP, 2019, pp. 456-460.
- [12] Paulus, J. and Virtanen, T.. Drum transcription with non-negative spectrogram factorisation. European Signal Processing Conference (EUSIPCO), 2005.
- [13] Dittmar, C., and Gärtner, D.. Real-Time Transcription and Separation of Drum Recordings Based on NMF Decomposition. International Conference on Digital Audio Effects (DAFX), 2014, pp. 187-194.
- [14] Wu, C. and Lerch, A.. Drum Transcription Using Partially Fixed Non-Negative Matrix Factorization with Template Adaptation. International Society for Music Information Retrieval Conference (ISMIR), 2015, pp. 257-263.
- [15] Vogl, R., Widmer, G., and Knees, P.. Towards multi-instrument drum transcription. Proceedings of the 21st International Conference on Digital Audio Effects (DAFx-18), 2018.
- [16] Böck, S. and Widmer, G.. Maximum filter vibrato suppression for onset detection. International Conference on Digital Audio Effects (DAFX), 2013.
- [17] Lee, H., Yoo, J., Choi, S.. Semi-Supervised Nonnegative Matrix Factorization. IEEE Signal Processing Letters, 2010, vol. 17, no. 1, pp. 4-7.
- [18] 瀧浪弘章, 若林友晴, 小白井敏明. 指数時間重み特性の影響を除去して騒音レベルのサンプリング値から時間平均音圧レベルを算出する方法. 騒音制御, 1999, vol. 23, no. 3, pp. 183-188.
- [19] 中村健太郎. 図解雑学 音のしくみ. ナツメ社, 2005.