

聴衆の知覚音量バランス推定と可視化によるドラム練習支援手法

細谷 美月[†] 中村 聡史[‡] 森勢 将雅[‡] 吉井 和佳^{†††}

[†] 明治大学大学院先端数理科学研究科 〒164-8525 東京都中野区中野 4-21-1

[‡] 明治大学総合数理学部 〒164-8525 東京都中野区中野 4-21-1

^{†††} 京都大学 〒606-8501 京都市左京区吉田本町

E-mail: [†]mitsumameggul8@gmail.com,

[‡] satoshi@snakamura.org, [‡]mmorise@meiji.ac.jp, ^{†††}yoshii@i.kyoto-u.ac.jp

あらまし 複数楽器を同時に演奏するドラムにおいて、主要な3楽器の音量バランスが重要である。しかし、聴衆とドラマーで聴こえ方が異なるため自身の音量バランスを把握できず、演奏を録音し確認する方法は時間や手間がかかるという問題があり、個人練習が難しい。我々はこれまで、ドラム正面に設置した単一指向性マイクから演奏を認識し、聴衆が知覚する音量を推定・可視化するシステムを提案および実装してきた。しかし、使用実験の結果、音量バランスの推定精度や結果の可視化方法などに問題があった。そこで本研究では、様々な音量の組み合わせやリズムで作成したドラム音源256パターンに対する音量バランス評価実験を行い、聴衆の知覚する音量バランス推定手法を提案する。また、システムの改善と評価実験の結果、本システムによって音量バランス練習を支援でき、人や状況によって結果の更新頻度を切り替えることで汎用性が高まる可能性が示唆された。

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

複数の楽器を同時に演奏することによってリズムや音量、音色などを表現するドラムは、“バンドの指揮者”とも呼ばれる。このドラムの上達に重要な技術の1つとして、メインの楽器となるハイハット (HI)、スネアドラム (SD)、バスドラム (BD) の3楽器の音量バランスがある (図1)。一般的に、BD・SD・HIの順に音量が大きく聴衆に知覚されるように演奏することが理想とされており、他の楽器より HI の音量が大きいと抑揚がない演奏に、BD の音量が小さいと安定感のない演奏に聴こえてしまう。しかし、ロックやメタルなどの激しい曲は SD と BD 両方を大きい音量に、ジャズは HI の音量を少し大きめに聴かせたいといったように、曲やジャンル、個人の好みによって目標とする音量バランスが異なる場合もある。このように、ドラマーは演奏をより良いものにするために、聴衆に自身の演奏をどう聴かせたいかによって目標とする3楽器の音量バランスを決め、それを意識しながら演奏することになる。

ここで、この音量バランス練習をドラマー個人で行うことは容易ではない。理由としては、叩く力が同じでも楽器により音量は異なるうえ、ドラマーが演奏しながら知覚する音量バランス (演奏中の聴覚と手や足の力加減の感覚などから判断する) と、ドラムから離れた位置にいる聴衆がそのドラム演奏を聴いて知覚する音量バランスが異なるためである。演奏を録音・確認し、再度演奏し直すといった練習方法もあるが、この方法は時間や手間がかかってしまう。また、ドラム演奏の練習支援に着目した既存研究は様々あるが



図1 ドラムにおけるメインの3楽器

[1,2,3], 音量バランスに着目した研究は少ない。

こうした問題をふまえ、我々はドラムの正面に設置した単一指向性のマイクから演奏を認識し、3楽器へ音源を分離した後に聴衆が知覚する音量を推定し、ドラマーに可視化して提示する仕組みを実現した[4]。しかし、音量バランスの推定精度が不十分であることや結果の可視化がわかりづらいといった問題があった。

そこで本研究では、様々な音量の組み合わせやリズムで作成した256パターンのドラム音源を用意し、データセット構築者は3楽器の音量バランスを評価してもらうことで聴衆の知覚音量バランスのデータセットを構築する。また、聴衆の知覚を加味した音量バランス推定手法を提案する。さらに、結果の可視化方法についても改善を行い、ドラマーが自力で理想的な音量バランスを容易かつ効率的に練習することを支援できるかどうか調査するためのシステム評価実験を行う。

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。触覚提示によって練習支援を行うものとして、Holland ら[1]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ブレスレット（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，Imada ら[2]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することにより練習を支援した研究として，Iwami ら[3]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，ドラム演奏における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”には着目していない．

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推定するために，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．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）[5]や，畳み込み非負値行列因子分解（NMFD）[6]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）[7]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[8]．

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り，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. 聴衆の知覚音量バランスデータセット構築

我々のこれまでの研究におけるプロトタイプシステム評価実験[4]より，音量バランス推定手法の精度が不十分であることがわかった．そこで，様々なリズム・音量バランスで作成されたドラム演奏の音源を聴いてもらい，その音量バランス（各楽器の大小関係）を評

価してもらうことによって，聴衆が知覚するドラム演奏の音量バランスのデータセット構築を行った．その後，音源分離から算出された各楽器の特徴量と得られた評価値が対応づいた回帰式を求めることによって，より聴衆の知覚に近い音量バランスの推定を行うことを目指す．

3.1. 評価音源データ作成

本実験では，聴衆がドラム演奏を聴いて判断する音量バランスを調査するため，HI・SD・BD について 4 段階の音量（本実験において p, mp, mf, f と表現するが，その音量は筆者らの主観によって選定したものである）を組み合わせ 64 パターンの音源を作成した．また，各楽器の音数の割合によって聴衆が判断する音量バランスも変化すると考えられるため，HI と BD の音数がそれぞれ少ない楽譜と多い楽譜を組み合わせ 4 種類のリズム（図 2）について，音量の組み合わせ 64 パターンの音源の作成を行い，全 256 パターンの音源の評価に用いる音源データとした．なお，音源のテンポ（演奏の速さ）は BPM120 とし，1 パターンにつき 16 小節分（32 秒）の長さで音源を作成した．

3.2. 音源データへの評価方法

様々なリズム・音量バランスで作成されたドラム演奏の音源を聴いてもらい，その音量バランスを評価してもらった．評価は，ドラムもしくは DTM（MIX）経験者である 10～20 代の 12 名に，静かな環境を保つことが可能な大学の音響室で行ってもらった．評価者の負担を考慮し，1 人につき全 256 パターンの半分の 128 パターン分の音源について評価を行ってもらった．

評価の手順を説明する．まず，初めの 5 分間で実験についての説明・指示を行った．具体的には，128 パターンの様々なドラム演奏の音源の音量バランスを評価してもらうこと，評価対象の音源は繰り返し聴取してよいこと，指定したタイミングと自身の判断したタイミングで休憩しながら進めることなどを伝えた．そして，練習としてランダムに選定した 4 音源分について，実際の流れで聴取・音量バランスの評価をしてもらった後，128 音源分の評価を開始してもらった．

なお，評価対象の音源は Excel のシート内に列挙されており，セルをクリックすることで再生できるようになっている．



図 2 評価で使った 4 種類のリズム

表 1 回答される評価値の一覧

HI	SD	BD	評価式
1	1	1	HI=SD=BD
1	1	2	BD>HI=SD
1	2	1	SD>HI=BD
1	2	2	SD=BD>HI
1	2	3	BD>SD>HI
1	3	2	SD>BD>HI
2	1	1	HI>SD=BD
2	1	2	HI=BD>SD
2	1	3	BD>HI>SD
2	2	1	HI=SD>BD
2	3	1	SD>HI>BD
3	1	2	HI>BD>SD
3	2	1	HI>SD>BD

表 2 作成時の各楽器の音量に応じた重み付け

音量	重み付けの値
p	0.25
mp	0.5
mf	0.75
f	1

評価は HI・SD・BD の 3 楽器で構成されるドラム演奏に対する音量バランスについて Excel シート上へ “音源内の 3 楽器について小さく聴こえる順番（音量が小さく聴こえる楽器ほど小さい数字となる）” を 1～3 の数字を打ち込む形で回答してもらった。評価値は，“＝”，“＜”，“＞” の 3 つの記号で表現できる大小関係の全 13 パターンに対応づけられる（表 1）。3 楽器分数字を打ち込むとその組み合わせに対応した評価式が表示されるようにし、間違いがないよう式を確認しながら評価を進めてもらった。

3.3. データセット構築処理

実験により、全 256 パターンの音源について、6 人分（1 人 128 パターンを 12 人分）の評価値のデータが得られた。ここでは、得られた評価を平均したものを、各音源に対する聴衆の知覚するドラム演奏の音量バランスの正解データと定義することとした。

具体的には、各音源について、各楽器に回答された数値について合計が 6 になるように正規化した後に平均値を算出し、作成時の各楽器の音量に応じて重み付け（表 2）を行ったものを、各楽器の音量の大きさを表す評価値（音量評価値）の正解とした。

このようにして算出した 256 パターンの音源における正解の評価式と、各楽器の評価値をまとめたものを正解のデータセットとし、各楽器について、音源から得られる特徴量からその楽器の音量の大きさを表す評価値（音量評価値）を算出可能とする回帰式を求める。これらの回帰式を用いて各楽器の音量評価値を求めることにより、聴衆がドラム演奏を聴いて判断する音量バランスを推定できると考える。

4. 聴衆の知覚を加味したドラム演奏の音量バランス推定手法

4.1. 回帰式の定義

3.3 節で構築したデータセットを用いて、ドラム演奏に対して聴衆が知覚する音量バランスの推定を行う。HI・SD・BD の各音源の特徴量として、各楽器それぞれについて、“RMS 値”，“閾値（音源内の最大振幅の 5 割）を超えた振幅の平均”，“スペクトル重心”，“ゼロ交差数”を選定した。演奏において、最低 1 音各楽器が演奏される最小の長さであるため、1 小節間の演奏について特徴量を求める。これらの特徴量から、各楽器の音量評価値を推定するための回帰式を、重回帰分析によって求める。音源内の最大振幅の 5 割を超えた振幅の平均を特徴量として採用した理由は、本研究で用いる NMF による音源分離は完全な分離ができず他の楽器の音声が分離結果に混ざってしまうため、振幅の値が高い場合には実際に叩いた音である可能性が高いと考えたためである。

このように音源分離の精度は完全ではないため、本研究では完璧に分離できた状態を想定し、HI・SD・BD の各楽器の音源のピュアデータから先述の 4 種類の特徴量を用いて、各楽器の音量評価値を推定するための回帰式を作成する。

$$\begin{aligned}
 \text{音量評価値}_{HI} &= c_1 \times RMS_{HI} + c_2 \times AmpT_{HI} - c_3 \times \\
 &\quad SpectralCentroids_{HI} + c_4 \times ZeroCrossing_{HI} + c_5 \\
 \text{音量評価値}_{SD} &= c_1 \times RMS_{SD} + c_2 \times AmpT_{SD} - c_3 \times \\
 &\quad SpectralCentroids_{SD} + c_4 \times ZeroCrossing_{SD} + c_5 \\
 \text{音量評価値}_{BD} &= c_1 \times RMS_{BD} + c_2 \times AmpT_{BD} + c_3 \times \\
 &\quad SpectralCentroids_{BD} - c_4 \times ZeroCrossing_{BD} - c_5
 \end{aligned}$$

256 パターンのドラム演奏音源の作成時に MIX を行う前の 3 楽器の音源から 4 種類の特徴量を算出し、3.3 節で構築した正解データセットと対応させたもので重回帰分析を行った結果、HI について $c_1=423.43070$, $c_2=2.52142$, $c_3=-0.00132$, $c_4=0.00002$, $c_5=6.32279$, SD について $c_1=117.05097$, $c_2=0.28139$, $c_3=-0.00607$, $c_4=0.00110$, $c_5=9.56647$, BD について $c_1=28.61916$, $c_2=6.85637$, $c_3=0.00037$, $c_4=0.00014$, $c_5=-0.67814$ となった。

各変数を正規化して重回帰分析を行った結果、偏回帰係数は表 3 のようになった。また、これについて K-

表 3 各楽器の回帰式の偏回帰係数

変数名	HI	SD	BD
RMS	2.30	2.08	0.78
閾値を超えた振幅	0.08	0.11	1.61
スペクトル重心	-0.45	-3.71	0.54
ゼロ交差数	0.06	3.79	-0.38
切片	1.26	1.41	1.43

表 4 回帰式の評価

評価指標	HI	SD	BD
MAE	0.20	0.16	0.29
MSE	0.07	0.04	0.14
RMSE	0.26	0.22	0.38
R2 スコア	0.87	0.94	0.79

分割交差検証を行い、平均絶対誤差 (MAE)、平均二乗誤差 (MSE)、二乗平均平方根誤差 (RMSE)、R2 スコアについては表 4 のような結果となった。

4.2. 推定手法の精度

これらの回帰式から各楽器の音量評価値を求め、その大小関係を表す式をドラム演奏から聴衆が知覚する音量バランスとする。音量バランス推定手法の精度については、3.3 節で構築した正解のデータセットにおける 256 パターンの音源について、2 楽器ずつの組み合わせ (HI と SD・SD と BD・BD と HI) の間に音量の差があると聴衆が知覚しているものについて、回帰式でも同様にその組み合わせを推定できたかによって判定した。

精度の算出方法を以下に示す。ここで 2 種類の楽器 A, B の音量評価値を $V(A), V(B)$ としたとき、 $|V(A), V(B)| < \delta$ である場合は A と B との間に差はないとみなし、そうでない場合は両者に差があるとみなす。そのうえで、正解データの評価値と音量バランス推定手法によって推定された評価値についてそれぞれの A と B の差を算出し、その差の値によって A と B の大小関係を求める。この δ を 0.3 とし、HI と SD, SD と BD, BD と HI の組み合わせそれぞれにおいて大小関係の一致度合いを求め、256 パターンの音源について平均したものを精度とする。なお、 $\delta=0.3$ としたのは、正解データにおいてどの 2 楽器の組み合わせについても、聴衆は評価値の差が 0.3 以上ある場合を差と知覚している傾向がみられたためである。

4.1 章で定義した回帰式についてピュアデータをテストデータとし精度を求めた結果、ドラム演奏から聴衆が知覚する音量バランスの推定精度が 87.4% となった。しかし、実際にシステムで用いる音源分離後の各楽器の音源データでテストした結果、47.0% と低い精度となった。そのためシステム化にあたり音源分離後の各楽器の音源データを用いて回帰式を再度定義する。

5. ドラム演奏の音量バランス練習システム

5.1. 提案システム

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 HI・SD・BD のみを抽出した音源から各音量を計算し、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 (図 3)。本システムは、演奏の音量バランスの推定をリアルタイムに行い、可視化して表示された結果を確認することができ、テンポ (演奏の速さ) についても設定可能である。

また、推定結果の更新頻度が異なる 2 種類のシステムを実装し、更新頻度が低いシステムをシステム A、更新頻度が高いシステムをシステム B とする。システム A は 1 小節間の演奏に対して録音し、音量バランスを推定、推定の計算が終わり次第結果を表示、結果が表示されてから 1 小節間分の秒数経過した段階で再び録音開始、といったように、いつの演奏に対する結果かわかりやすくすることをねらいとした。また、システム B は、1 小節間の演奏に対して録音、録音した演奏について音量バランスを推定、推定の計算が終わり次第結果を表示 (2 拍分の秒数の間、画面が静止する) という処理を 1 小節ずつシフトしていくことによって、結果が常に更新され、よりリアルタイム性を高くすることをねらいとした。

5.2. システムにおける音量バランス推定手法

システム化にあたって、音源分離後の各楽器の音源データを正解データとして回帰式を再度算出した。また、リアルタイム性を保つためにはできるだけ計算量が少ない方がよいと考えたため、影響が少ないかつ計算に時間がかかる特徴量について削除を行った。システムにおける各楽器の音量評価値を推定するための回帰式を以下に示す。

$$\text{音量評価値}_{HI} = 843.896 \times RMS_{HI} - 0.149$$

$$\text{音量評価値}_{SD} = 349.734 \times RMS_{SD}$$

$$+ 0.001 \times \text{SpectralCentroids}_{SD} - 2.899$$

$$\text{音量評価値}_{BD} = 70.811 \times RMS_{BD} + 10.930 \times \text{AmpT}_{BD} - 0.088$$

5.3. 実装

システムは、プロトタイプシステムと同様、音声処理部分を Python、可視化や UI 部分を Processing で実

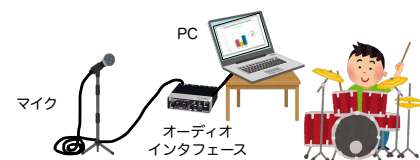


図 3 システム使用想定環境

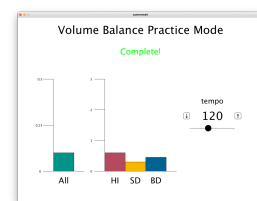


図 4 システム画面のスクリーンショット

装し、機材については、PC は MacBook Pro (13-inch)、マイクはダイナミックマイクの SM58、オーディオインターフェースは UR12 を使用した。

ユーザに提示するシステム画面を図 4 に示す。音量バランス可視化ゾーンにおける赤、黄、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 HI・SD・BD の音量の大きさを表す。グラフの縦軸は、その楽器の音量の大きさを示す値となっている。なお、画面右下のテンポ (BPM) 設定 UI により、練習するテンポを調整可能とする。バー上の丸をドラッグ、バー上の任意の点位置をクリックもしくは ↓・↑ ボタンをクリックすることによってテンポを調節可能である。

5.4. 利用方法

ユーザ (ドラマー) は、図 3 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、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、推定された聴衆の知覚する音量バランスが可視化されたシステム画面 (図 4) を見ながら演奏を変化させる。なお、図は入力されたドラム演奏について聴衆は HI>BD>SD の音量バランスで知覚するという推定結果を可視化した状態のシステム画面である。この図において、もしユーザが BD>SD>HI で聴衆が知覚するような演奏を目指している場合は、ユーザが HI の音量を小さく叩くように演奏を修正することで、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。

6. プロトタイプシステムを用いた実証実験

提案システムによって、ドラマーが自力で理想的な音量バランスを容易かつ効率的に練習することを支援できるかどうか調査するため、システムを使用しながらアコースティックドラムで目標の音量バランスに向けて演奏してもらう実験を行う。具体的には、指定した音量バランスを目指して演奏するタスクを複数回を行い、タスク終了後に自身の演奏の録音データを聴いてもらったうえで、システムに関するアンケートに回答してもらう。また、推定結果の更新頻度が低いシステム (システム A)、更新頻度が高いシステム (システム B)、システムなしの 3 条件の練習環境条件で音量バランス練習を行ってもらい、練習環境の違いによる音量バランス練習のしやすさについても明らかにする。

音量バランス推定の精度、可視化方法のわかりやすさ、音量バランス練習のしやすさについて、アンケートへの回答を用いて評価を行う。システムによってドラマーが自力で理想的な音量バランスを容易かつ効率的に練習可能となるか検証するために、本実験では「システムなし条件に比べてシステムあり条件の方が、評

価指標の値が高くなる」という仮説を立て、分析する。

6.1. 実験手順

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 6 名である。実験の様子を図 5 に示す。

実験では、まず事前準備として、システムと機材 (PC、マイク、オーディオインターフェース) のセッティングを筆者らが行った。図 3 のようにドラムから見やすい位置に PC、ドラムから 2m 離れた位置かつ床から約 1m60cm の高さにマイクを設置し、PC は MacBook Pro (13-inch, 3.5 GHz デュアルコア Intel Core i7, 16GB)、マイクは単一指向性マイクである Shure の SM58、オーディオインターフェースは Steinberg の UR12 を使用した。そして、実験協力者にドラムのセッティングと、ドラムに慣れるための自由演奏を行ってもらった後、実験に関する指示とシステムの説明を行った。具体的には、指定した条件・順番で目標の音量バランスを目指して演奏すること、目標の音量バランスに到達できたと判断したタイミングから、4 小節間その音量バランスを保ったまま演奏を行った後すぐ演奏を終了すること、演奏開始から 2 分間 (60 小節) 経過した時点でその条件については「到達不可」と判断すること、実験中の演奏は録音を行うことなどを伝えた。また演奏の制約として、テンポは BPM120、使用する楽器は使用する楽器は HI・SD・BD の 3 楽器のみで、HI はオープン奏法を使用せずクローズ奏法のみで演奏すること、SD はリムショット奏法を使用しないことを伝えた。

指示と説明が終了した後、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実験協力者に行ってもらった。実験では、実験協力者 1 人につき 5 パターン×練習環境 3 条件の全 15 パターンを行ってもらった。タスクにおいて指定した演奏条件は、一般的に目標とされやすい音量バランス 3 条件 (BD>SD>HI, SD>BD>HI, HI>BD>SD) と、音量バランス推定手法を定義する際に使用した 4 種類のリズムの組み合わせの 12 パターンから、16 ビートかつ「BD>HI>SD」である 2 パターンを除外した全 10 パターンとした。この 2 パターンを除外した理由は、プレ実験において「16 ビートにおいて SD を小さく演奏することがそもそも難しかった」という意見がみられたためである。実験では、システム A・システム B・システムなしの練習環境条件それぞれについてランダムに選定した 4 パターンの演奏を行ってもらった。また、音量バランス推定手法が様々なリズムに対応可能か検証するために、各練習環境条件において、実験協力者に 4 種類のリズム以外の自由なリズム、3 条件の中から習得したい音量バランスで目標を設定してもらって演奏する 1 パターンを行ってもらった。

表 5 実験で使ったアンケート項目一覧

	内容
Q1	システム or 自分は聴衆が知覚する音量バランスを推定できていたと思うか
Q2	自身の演奏の音量バランスの可視化を見ながら演奏することで音量バランス練習はしやすかったか
Q3	結果が切り替わる頻度は適切だったか
Q4	結果表示方法はわかりやすかったか
Q5	音量バランスを変化させたときシステムにその結果が反映されていた感じがしたか
Q6	今後このシステムを使って練習したいか
Q7	自分の音量バランスに自信をもって練習できたか
Q8	システム A・B・システムなしについて練習しやすかった順に 1～3 の番号を回答してください
Q9	システムや実験についての感想

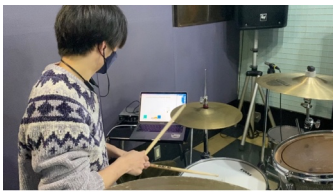


図 5 実験の様子

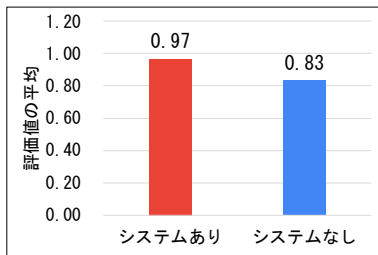


図 6 Q1 へ回答された評価値の平均

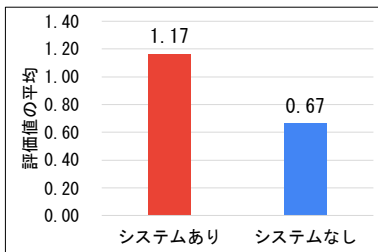


図 7 Q7 へ回答された評価値の平均

タスク中の全ての演奏について録音・録画を行い、システム使用時は音量バランスの推定に用いた 4 種類の特徴量（RMS 値・閾値（音源内の最大振幅の 5 割）を超えた振幅の平均・スペクトル重心・ゼロ交差数）と、演奏開始からの経過時間のデータについても取得を行った。またタスク終了後、タスク中の演奏の終盤の録音を聴いてもらい、システムの使いやすさや今後使いたいといった内容についてのアンケートやインタビューを行い、実験終了とした。具体的なアンケート内容を表 5 に示す。

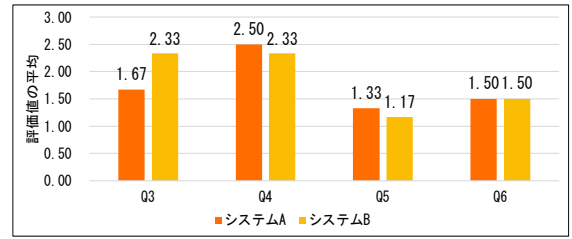


図 8 Q3～Q6 へ回答された評価値の平均

表 6 Q8 への回答

回答	人数
A>B>なし	2 人
B>A>なし	3 人
なし>B>A	1 人

6.2. 結果

システムあり条件とシステムなし条件の Q1「システム or 自分は聴衆が知覚する音量バランスを推定できていたと思うか」へ回答された評価値の平均を図 6 に示す。この図より、システムあり条件の方がシステムなし条件より精度が良い傾向がみられた。Q2「自身の演奏の音量バランスの可視化を見ながら演奏することで、音量バランスの練習はしやすかったか」へ回答された評価値について、平均は 1.83, 分散は 0.47 となり、どの実験協力者も+の評価をしていた。次に、Q7「自分の音量バランスに自信をもって練習できたか」へ回答された評価値の平均を、練習環境条件ごと（システムあり・システムなし）にまとめたものを図 7 に示す。この図より、システムあり条件の方がシステムなし条件より自分の演奏の音量バランスに自信をもって練習できていることがわかる。

次に、システム A 条件とシステム B 条件の Q3「システムの結果が切り替わる頻度は適切だったか」、Q4「システムの結果表示方法はわかりやすかったか」、Q5「音量バランスを変化させたときシステムにその結果が反映されていた感じがしたか」、Q6「今後このシステムを使って練習したいか」へ回答された評価値の平均を図 8 に示す。この図より、Q3「システムの結果が切り替わる頻度は適切だったか」については、結果の更新頻度の高いシステム B が高い評価となり、Q4「システムの結果表示方法はわかりやすかったか」、Q5「音量バランスを変化させたときシステムにその結果が反映されていた感じがしたか」、Q6「今後このシステムを使って練習したいか」への回答については、システム A とシステム B に大きな差はなかった。

Q8「システム A・システム B・システムなしについて練習しやすかった順に 1～3 の番号を回答してください」への回答を、表 6 に示す。この表より、実験協力者ごとに回答は様々であることがわかる。

7. 考察

7.1. 提案手法による音量バランス練習支援効果

「システムなし条件に比べてシステムあり条件の方が、各評価指標の値が高くなる」という仮説をもとに、システムによってドラマーが自力で理想的な音量バランスを容易かつ効率的に練習可能となっているか検証する。

図7より、システムあり条件の方がなし条件より自分の演奏の音量バランスに自信をもって練習できていることがわかる。実際に、システムなし条件について「完全に自分の感覚で叩いていたので、不安が多かった」といった回答がみられた。聴衆が知覚する音量バランスを可視化することによって、納得感をもって音量バランス練習を行うことができると考えられる。

また、Q2について全ての実験協力者が+の評価であり、表6について6名中5名がシステムなしよりシステムありの評価が高いことから、提案システムによって音量バランス練習がしやすくなっているとわかる。

これらの結果について、「システムなし条件に比べてシステムあり条件の方が、評価指標の値が高くなる」という仮説通りの結果となり、本システムによってドラマーが自力で理想的な音量バランスを容易かつ効率的に練習することを支援できる可能性が示唆された。

7.2. 推定結果の更新頻度による練習しやすさ

結果の更新頻度の適切さについて、図8のQ3より、頻度の高いシステムBの方が高い評価となり、実際に「リアルタイムで反映された方が使いやすい」「即時反映されるので、練習に使うならこちらが使いやすい」といった意見もみられた。一方、頻度の低いシステムAについて「音量バランスの安定感を測るのに良いシステム」「部分的にフレーズを練習するときに使える」といった意見がみられた。このことから、結果の更新頻度を変化させることで、自分の演奏の音量バランスを細かく調整したい場合、曲の一部分を練習したい場合など様々な練習に対応できる可能性が示唆された。

結果表示方法のわかりやすさについてはQ4より、システムAとシステムBの評価に大きな差はなかった。実際に「Recordingのタイミングが分かりやすく、練習に反映しやすく感じた」「いつ演奏した分の結果がわかりやすく、やりやすかった」といった意見がみられた。

また、Q5より、音量バランスを変化させたときシステムにその結果が反映されていた感じがしたかについても、システムAとシステムBの評価に大きな差はなかった。結果の更新頻度が低いシステムAについて高い評価となった。これについては、システムによる音量バランス推定の精度の高さが不十分であったことか

ら、更新頻度の高いシステムBについて結果がより反映されていないと感じやすかったのではないかと考えられる。

今後このシステムを使って練習したいかについて、Q6よりシステムAとシステムBに大きな差はなかった。またQ8より、システムA・システムB・システムなしについて練習しやすかった順番を回答してもらった結果は、実験協力者ごとに回答は様々であった。さらに、Q3～Q6についても実験協力者ごとに回答が様々であり、人や状況によって音量バランスの結果の更新頻度を切り替えることで、より音量バランス練習を支援できると考えられる。例えば、結果の更新頻度については、音量バランスを一定に保つといったように安定感を鍛えたい場合や部分的なフレーズを練習したい場合は更新頻度を少なくし、目標とする音量バランスを目指して演奏を調整したい場合は更新頻度を高くするといったことが考えられる。

7.3. 聴衆の知覚する音量バランス推定精度

実際に、「自分が聴きながら感じるバランスと、聴こえるバランスは違うのだと思った」「HIの音量が録音してみると以外と大きく聴こえたりと、自分の想定より力が入ってしまっているのに気づくという上でこのシステムは有用だなと感じた」といった意見が得られたため、ドラマーと聴衆の知覚の差によって自身の演奏の音量バランスを把握しづらい問題をこのシステムで解決できると考えられる。

しかし、Q9「システムや実験についての感想」への回答のうち、システムの音量バランス推定精度に関するものについて、主に精度が十分ではないといった意見が多くみられた。その中でも特に、「SDは小さい音でも大きく結果が出てしまう感じがあった」「ほぼSDを叩いてないのに、音量が出ている結果になった」といったようにSDの音量の推定精度に関する回答が多かった。この問題については、SDとBDの周波数帯域が似ており、音源分離の際にSDの音源にBDの叩打音が混ざってしまうために競合してしまったと考えられるため、本研究で採用したNMFやSSNMFについてより精度が高められる学習回数をみつけることや、他の音源分離手法を検討していくことで、解決できると考えられる。

また「単純な音量だけでなく音色によって可視化された音量が変化していたので、逆にわからなくなる部分があった」「倍音多めの音は、SDの音量が下がっている気がした」といった意見から、叩く際の音色によって出力される音量に違いがある可能性が示唆された。本システムでは、各楽器について筆者が演奏した叩打音1音のみを基底としてSSNMFによる音源分離を行ったが、打楽器は叩く位置・角度・勢い・強さなどに

よって音色が様々変化する．今後は様々なパターンの基底を用意することで，より汎用性の高い音量バランス推定手法が実現可能であると考ええる．

7.4. 制約と展望

音量バランス推定手法について，今後様々なリズムパターンについて聴衆が知覚する音量バランスのデータを増やすことや，リズムの違いによる各楽器の音数の割合の差が聴衆の判断する音量バランスにどう影響を及ぼすのかについてより詳しく調査することによって，より幅広いリズムパターンを対応可能なシステムにしていく予定である．また，現状対応できていないHIのオープン奏法やSDのリムショット奏法のように音色が異なる奏法について基底を作成して音源分離を行い，音量が大きいものを採用するといった方法によって，対応可能であると考えられる．

聴衆が知覚するドラム演奏の音量バランスは，演奏が行われる場所や音色によって異なるため，今後は演奏環境の変化への対応についても検討する．また精度については，音源分離の精度を上げ3楽器を綺麗に分離可能となれば向上すると考えられる．今後は本研究で採用したSSNMFについてより精度が高められる学習回数をみつけることや他の音源分離手法の検討によって，精度を高めていく．

システムについて，音量バランス推定結果の適切な可視化方法は使用する状況によって変化することや人による好みがあることがわかった．また，音量バランスの表現方法について，「全体音量との比較ではなく3楽器の比率で見られると見やすいかも」「人が生音で演奏を聴いた時，ライブハウス等マイクを立ててスピーカーで音を出す時，レコーディング時のバランスなど，いくつかのモードがあったらもっと面白くなると感じた」といったような意見がみられた．今後は練習したい状況によって結果の更新頻度や表示方法を切り替えるといった機能を検討していく予定である．

8. おわりに

本研究では，ドラム演奏の音量バランス練習をより容易で効率的なものにすることを目的とし，マイクから入力したドラム演奏から聴衆が知覚する音量バランスを推定し，結果を可視化して提示するシステムを提案した．これまでの研究でプロトタイプシステムを実装し評価実験を行った結果，音量バランス推定手法や可視化手法に問題があった．そこで，様々なドラム演奏の音源についてその音量バランスを評価してもらうことで聴衆の知覚音量バランスの正解データセット構築を行い，得られた評価データを用いて，ドラム演奏の録音データから聴衆が知覚する音量バランスの算出をピュアデータについて87.4%の精度で可能とする音

量バランス推定手法を提案した．改善したシステムを用いて実験を行い，その有用性について評価を行った結果，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提案手法によってドラマーが自力で理想的な音量バランスを容易かつ効率的に練習することを支援できる可能性や，人や状況によって音量バランスの結果の更新頻度を切り替えることで，より提案手法によって音量バランス練習を支援できる可能性がみられた．

今後は，音量バランス推定手法精度の向上を行うことで，音量バランス練習をより支援できると考える．また，バンド，アンサンブルやオーケストラでの演奏などにおける各楽器パートの音量バランス練習にも応用する予定である．

参 考 文 献

- [1] Holland, S., Bouwer, A. J., Dalgelish, M. and Hurtig, T. M.. Feeling the beat where it counts: fostering multi-limb rhythm skills with the haptic drum kit. The 4th International Conference on Tangible and Embedded Interaction, 2010, pp. 21–28.
- [2] Imada, Y. and Ochi, Y.. Hitting Arm Detection for Drum Performance Learning System Using Kinect. The Seventh International Conference on Informatics and Applications (ICIA2018), 2018, pp. 8–10.
- [3] Iwami, N. and Miura, M.. A Support System for Basic Practice of Playing the Drums. International Computer Music Conference, 2007, pp. 364–367.
- [4] 細谷美月, 中村聡史, 森勢将雅, 吉井和佳. 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, 情報処理学会 研究報告エンタテインメントコンピューティング (EC), Vol.2021-EC-59, No. 27, pp. 1–8, 2021.
- [5] Lindsay-Smith, H., McDonalds, S. and Sandler, M.. Drumkit transcription via convolutive NMF, 15th International Conference on Digital Audio Effects (DAFx), 2012.
- [6] 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.
- [7] Vogl, R., Dorfer, M. and Knees, P.. Recurrent Neural Networks for Drum Transcription. The 17th International Society for Music Information Retrieval Conference (ISMIR), 2016, pp. 730–736.
- [8] Wu, C. W., Dittmar, C., Southall, C., Vogl, R., Widmer, G., Hockman, J., Müller, M. and Lerch, A.. A Review of Automatic Drum Transcription. IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 26, No. 9, pp. 1457–1483.