

# 音響情報と動画像情報を用いたドラムソロ演奏の採譜法の改良

村松 慧一 野口 将人 坪井 邦明 松島 俊明  
 東邦大学 理学部 情報科学科

## 1. はじめに

自動採譜の研究は数多く行われているが、打楽器を対象としたものは少ない。ドラムの自動採譜を試みた研究例もあるが、噪音であるドラム音を周波数分析により識別することは難しい[1][2]。そこで、音響情報に加え、発音ピーク時の画像からスティック頭部の位置を検出し、発音候補を絞ってから音響情報による自動採譜を試みたが、十分な絞り込みを行うことが出来なかった[3]。今回、発音ピーク時前後のスティック頭部の動きと、発音ピーク時以降の周波数帯域毎の減衰の様子を用いた所、発音候補楽器の絞り込み精度が大幅に向上した。今回報告するシステム全体の処理の流れを図1に示す。「スティックの動作解析による楽器候補の選出」と「数フレーム後の音響情報の処理」が今回新たに修正・追加した処理である。



図1. 全体の処理の流れ

## 2. ドラムスティックの動作解析

ドラムスティックの動作解析のために図2のようにあらかじめドラムセット領域の設定を行う。以前報告した



図2. ドラムセット領域の設定



図3. ドラムスティックの動作解析

方法[3](以後「従来法」と呼ぶ)では、ドラムセットの上部から撮影していたが、本研究ではドラムスティックの動きが分かり易い横方向から撮影した動画像を用いた。また、スティック頭部にはそれぞれ赤、青の2色のマーカーをつけた。

まず、音響信号の発音ピーク時刻の画像フレームを抽出し、スティック上のマーカーを検出して発音ピーク時刻のスティック頭部の位置を決定する。その後発音ピーク時刻の前と後の画像フレームを抽出し、それらの画像内のマーカー位置からマーカーの移動ベクトルを算出し、スティックの動作解析を行う。例えば図3のように、発音ピーク時刻にマーカーがドラムセット領域内にあっても、青色マーカーのようなスティックの動きは実際には楽器を叩いていないと判定する。

## 3. 音響情報処理

ドラム音は噪音のため、各楽器のピッチが正確に決まってはいるが、特定の周波数帯域のパワースペクトルの発音後150msまでの減衰の様子を調べた結果、緩やかに減衰する膜鳴楽器のスネアドラム(SD: 周波数帯域 180-200Hz)、急激に減衰するハイタム(HT: 200-250Hz)、ロータム(LT: 100-130Hz)、フロアタム(FT: 130-160Hz)、バスドラム(BD: 130-160Hz)のSD以外の膜鳴楽器群、ほとんど減衰しないハイハットシンバル(HH: 450-500Hz)、クラッシュシンバル(CC: 300-400Hz)、ライドシンバル(RC: 400-430Hz)の体鳴楽器群の3つに大きく分類できることが分かった。図4に各楽器群の特定の周波数帯域毎の減衰例を示す。この特徴の違いに注目することで実際に発音している楽器群の候補の絞り込みを行う。

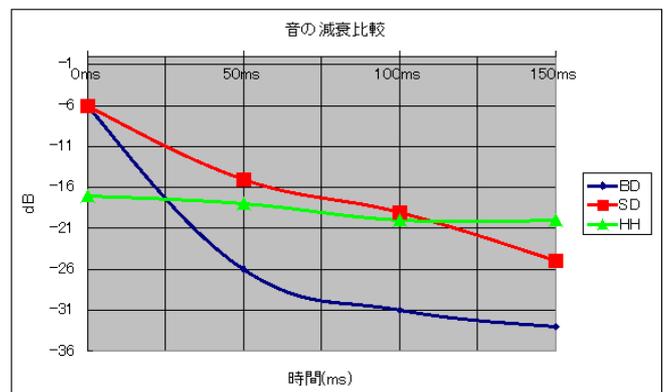


図4. 特定の周波数帯域のパワー減衰例

Improvement of the drum solo music transcription using the acoustic and video information  
 Keiichi Muramatsu, Masato Noguchi, Kuniharu Tsuboi and Toshiaki Matsushima  
 Dept. of Information Science, Faculty of Science, Toho University

#### 4. 楽器判別方法

基本的にドラムは2本のスティックと右足で踏むBDで演奏される。従って発音している楽器の候補の組み合わせは、何も演奏されていない場合も含めると右手、左手、右足の組み合わせの最大8通りである。まずスティックの動作解析を行い、右手または左手の楽器で叩いていないと判断できる楽器を含む候補を除外する。スティックの動作解析で叩いていると判断した場合でも実際には叩いていない場合があったり、BDが叩かれたどうかは画像解析では判定できないため、特定の周波数帯域のパワーの減衰の様子を調べて実際にその楽器が鳴っているかどうかの判定を行う。以上の絞り込みの結果で残った候補楽器群について、従来法で用いていた周波数パターン解析による楽器認識を行い、最終的な楽器の特定を行う。

#### 5. 実験結果

本手法の有効性を確認するために、まず、BD以外の楽器について、スティックの動作解析の認識精度を調べる実験を行った。その実験結果を表1に示す。この結果から、スティックの動作解析が叩いている楽器の絞り込みに有効であることが分かった。

表1. スティックの動作解析による楽器判定結果

楽器	正解数/全数	楽器	正解数/全数
HH	10/10	HT	9/10
CC	10/10	LT	10/10
RC	8/10	FT	9/10
SD	8/10		

次に、異なる演奏速度による実際のドラム演奏の画像と音響データを用いて、楽器判別の実験を行った。誤判定の楽器を赤色の音符で示した実験結果を図5～図7に示す。ただし、発音ピーク時およびその前後の画像・音響データの切り出し、採譜結果の楽譜化は手動で行った。



図5. 採譜結果 (BPM=80)



図6. 採譜結果 (BPM=100)



図7. 採譜結果 (BPM=120)

従来法では非常に誤りが多かったHHとSDの組み合わせに対して、本手法では大幅な改善が見られた。また特定の周波数帯域の減衰の様子を調べた結果、異なる楽器群の組み合わせ候補に対しては従来法と比較すると精度の向上が見られたが、同じ楽器群に属する楽器同士の組み合わせ候補についてはあまり改善が見られず、実際には演奏されていない楽器が演奏されていると判定する誤検出が生じた。また、BPM=80では3ヶ所、BPM=100では4ヶ所、BPM=120では6ヶ所と、同じ楽曲でもテンポが速くなるほど誤検出が多くなることが分かった。

#### 6. まとめ

実験結果から、HHとSDの組み合わせにおいて誤検出が少なくなり、結果的に楽曲全体に対する判定精度が向上した。一般的なドラム演奏においてもHHを叩きながらSDを叩くパターンが頻出するため、実験に使用した楽曲以外に対しても本手法が効果的であると予想できる。また、特定の周波数帯域のパワーの減衰の様子を調べる方法は、異なる楽器群の組み合わせ候補については効果的であるが、同じ楽器群に属する楽器では特定の周波数帯域の減衰の様子に大きな違いがないため楽器候補を絞ることが難しいことが分かった。また、テンポが速くなると、減衰の様子を調べる対象である150msの範囲内に次の発音ピークが出現する場合が生じるため、誤検出が多くなると考えられる。今後は様々な楽曲や演奏に対し実験を行うと共に、データ切り出しと結果の楽譜化の自動化を行っていく必要がある。

#### 文献

- [1] 吉井和佳, 北原鉄郎, 桜庭洋平, 奥乃博: 自己組織化マップによる教師なしクラスタリングを利用したドラム演奏の自動採譜, 情処研報 MUS-51, 43-50 (2003)
- [2] 田中徹, 田中敏幸: ドラム音の種類判別及び自動採譜, パーソナルコンピュータ利用技術学会全国大会講演論文集, 3, 113~116 (2008)
- [3] 大矢健二郎, 松島俊明: 音響情報と動画像情報を用いたドラム演奏の自動採譜, 信学総大 A-10-12 (2013)