

# センサデータと機械学習を用いた ドラム演奏動作認識手法の検討

## Drum Performance Recognition Method Using Sensor Data and Machine Learning

今田 泰広<sup>1</sup> 越智 洋司<sup>2</sup> 井口 信和<sup>2</sup>

Yasuhiro Imada<sup>1</sup> Youji Ochi<sup>2</sup> and Nobukazu Iguchi<sup>2</sup>

<sup>1</sup> 近畿大学大学院総合理工学研究科

<sup>1</sup> Graduate School of Science and Engineering, Kindai University

<sup>2</sup> 近畿大学理工学部

<sup>2</sup> Faculty of Science and Engineering, Kindai University

**Abstract:** Stable drum playing requires learning of appropriate tapping and usage of stick to a player. In this study, we developed a proto-type system that can analyze the information of various sensors and objectively evaluate the player's performance. This system realized arm detection and play style identification by focusing on the joints of arms. In this paper, we consider a motion recognition method using machine learning technology and its usefulness from the evaluation results.

### 1. はじめに

音楽演奏において安定したリズムの習得にはそれらに対する適切な奏法の習得とその使い分けが必要である。そのための練習においては、音に注目するだけでなく身体動作にも注目する必要がある。西村ら[1]の研究ではドラム演奏の熟達者と初心者の演奏動作の違いは体格よりも奏法に現れることが示唆されており、演奏動作は重要な要素となる。

練習は一人で行う場合と、指導者のもとで行う場合がある。指導者の監督のもとであれば、演奏者が正しい動作で演奏できているかを客観的に確認できる。しかし、常時指導を見ることはできないため、一人で練習する時間がある。その場合で客観的に動作を確認するためには、鏡を使う、動画を撮って確認するなどの方法がある。ただし、奥行き動作などこれらの方法で確認できない情報や、練習者が想定していない動きまで確認することは難しい。

このような問題を解決するために ICT を用いた練習を支援する研究は、音楽演奏だけでなく芸術、スポーツ、医療など運動においてトレーニングを必要とする分野で広く行われている。

本研究ではドラムセット（以下、ドラム）を使用した楽器演奏を対象とする。我々は、ドラム演奏を対象とする学習支援システムの研究を行ってきた。その一環として過去にモーションキャプチャセンサの一つである Kinect で練習者のドラム演奏姿勢を検出し、自主練習時にも演奏フォームを判断できるドラム練習支援システム D-Learning[2]を開発した。

このシステムでは両腕の肩・肘・手首の関節に対して教則データと比較し譜面と対応させて類似度を確認でき、類似度に応じてコメントが表示できる。しかし、あくまで腕の関節の動きが教師データと比べてどの程度異なるかがわかるだけで、実際に叩いた楽器の種類や、リズム、叩く強さ、音色などは判定できなかった。すなわち、身体動作と実際の演奏結果の関連性についてまで指摘できなかった。そのため、楽譜と異なった演奏を行なった場合に学習者自身が認識して修正する必要がある。

そこで、我々は叩いたタイミングの認識と、そのときどのような身体動作であったかを識別する手法を提案してきた。叩いたタイミングに関しては、電子ドラムを利用して MIDI メッセージを取得することで認識している。どのような動作で叩いたかに関しては、Kinect から取得した距離情報と、RGB 画像から右手と左手どちらの腕で叩いたかを識別する手法（以下、叩打腕識別手法）[3]、腕の振りの大きさを識別する手法[4]を提案してきた。叩打腕識別手法では、ドラム打面とドラムスティックのインタラクションを認識する必要があるため、それぞれの位置情報が必要となる。ドラムスティックの位置に関しては、カラーマーカーを利用することで特別なキャリブレーションなど必要なく推定することができた。しかし、ドラム打面の位置を登録するには目視で位置を確認し指定する必要があった。さらに、識別ルールも閾値の設定などヒューリスティックな方法で設定していたため、ドラムやカメラの位置が変わるたびに再設定し直す必要があった。

そのため、本稿では機械学習手法を用いて自動的に腕やドラムスティックなどの位置を学習し、叩打腕を識別する手法を提案し、その妥当性を検証する。時系列データなどに利用される LSTM を利用し、叩いた瞬間から 15 フレーム前までのドラムスティックの位置などを学習させた結果、テストデータで約 9 割の正解率で識別することができた。本稿では機械学習手法の検証だけでなく、ドラム演奏のスキル学習支援のアプローチについても述べる。

## 2. ドラム奏法

ドラムは両手両足を使い、複数の打楽器を対象として演奏する。一般的に直接手で叩打するわけではなく、木製のドラムスティックを介して打面を叩打する。ドラムスティックをどのように制御するかは打楽器演奏において重要な要素である。ドラムはポピュラー音楽では「バスドラム」、「スネア」、「ハイハット」という 3 種類の楽器を中心に演奏する。他に複数のタムタム、シンバル等を配置する。バスドラムは右足でペダルを使って演奏する。ハイハットはスティックとペダルを両方使用する。その他の楽器はスティックを使用して演奏する。

音楽演奏を向上させるような学習はスキル学習として定義できる[5]。学習の要素として目標となる成果物の認識と行動のインタラクションが大きな要素である。特定の楽曲を練習する場合に、どの打面をどのようなリズムで叩打していくか（要素 1）を理解し、記憶する段階は認識の段階である。認識した楽曲を目標として、演奏者自身の身体をどのように動かすかは行動の学習である。実際に体を動かしながら、演奏された曲、動作の様子をフィードバックし習得する。

動作としてフィードバックできる要素は様々な切り口がある。たとえば左右のスティックをどのような順番で楽器間を移動させ叩打していくか（要素 2）などを考慮する必要がある。ドラムスティックのコントロールの方法でも奏法を分類できる。たとえば腕のどの関節を支点にするか（要素 3）で分類できる。支点をコントロールすることで、音量とリズムを調整できる。

## 3. ドラム奏法の練習支援モデル

前章で述べたフィードバック要素を初心者が自己確認することは困難だと考える。そこで、これらの要素を識別し診断するシステムが実装できれば、ドラム演奏の学習を促進する効果が期待できる。

本研究では主に一人で練習する場合で、自宅などの練習を考慮し電子ドラムの利用を前提とする。動作のフィードバックの支援対象は上半身とする（脚の動きは除外する）。これは、Kinect などのカメラを利用する場合に上半身と下半身で言えば上半身の動きのほうが撮影しやすいためである。よって、図 1 のようにドラム演奏者の前方上部に Kinect を設置する。D-Learning ではこのような環境でドラム奏法を識別し、熟達者の奏法と比較することで、自身の演奏姿勢の判断を支援できるシステムを構築した。今後もこれを踏襲し、さらに演奏結果についてもフィードバックすることを目的とする。すなわち、演奏結果と動作の対応付けを促進するシステムを構築することが本研究の目的となる。

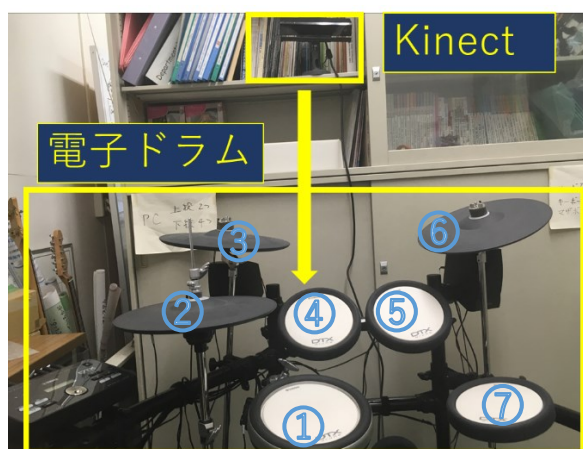


図 1: 想定環境

## 4. ドラム奏法の識別アプローチ

本稿では叩打腕の識別アプローチとして機械学習を採用した。その話をメインとし、その他のフィードバック要素の識別アプローチについても簡単に説明する。

### 4.1 演奏結果の識別(要素 1 への対応)

電子ドラムの MIDI メッセージから叩打タイミングと叩打した楽器の種類、叩打した強さなどを取得する。ドラムスティックがドラムヘッドに 1 回接触したときにこれらの情報が 1 まとまりの情報として取得できる。これを 1 ショットとすれば 1 ショットごとにどのような動作で叩打したかを識別することができる。以下に述べる叩打腕の識別、叩打支点の識別はいずれも 1 ショット叩打されたと判定されたときに行われる。

## 4.2 叩打腕の識別(要素 2 への対応)

1 ショット叩打されたときに左右どちらの腕で叩打されたのかを画像により識別するには、左右どちらのドラムスティックがどのドラムヘッドにどのタイミングで衝突したのかを認識させる必要がある。ドラムスティックとドラムヘッドの位置情報を推定させることでその位置関係や移動ベクトルから叩打腕を識別できる。以下、ドラムスティックとドラムヘッドの位置推定手法、それらの情報から叩打腕を識別する手法について述べる。

### 4.2.1 ドラムスティックの位置推定

ドラムスティックの先端に赤と緑のカラーマーカーを装着し、画像処理技術(OpenCV)を使用して左右のドラムスティックの位置を識別・推定する。背景を除去しノイズを軽減するために、Kinect から見て腰(SPINE\_BASE)より奥の画素をフィルタする。HSV 表色系に画像を変換し右は赤、左は緑の画素を抽出する。その画素に対応する 3 次元座標を Kinect の SDK の変換関数で算出する。以上の処理過程を経てドラムスティック位置を推定する。

### 4.2.2 ドラムヘッドの位置推定

以前報告した方法では、3 次元位置からドラムヘッドの画素領域を手動で抽出する方法でドラムヘッドの位置を推定していた。しかし、この方法ではカメラとドラムの位置関係が変わってしまうたびに抽出する必要がある。そこで、本稿ではドラムスティックの位置推定から推定した位置と叩打したドラムヘッドの対応関係に規則性があると考えて機械学習で自動的に学習することを期待する。そのため、特別にドラムヘッドの位置推定は行わない。

### 4.2.3 叩打腕識別手法

本稿では機械学習の一種の RNN (Recurrent Neural Network) を利用する。RNN では時系列データから特徴量を自動で学習することができる。そのため、例えばドラムスティックの軌跡のパターンを抽出することを期待し採用する。

学習モデルの入力をドラムスティックの推定位置、腕の関節角度、演奏結果等とする。これらを叩打の瞬間から一定以前のフレーム分取得する。そして、出力を右か左を表すラベルとする。この形式のデータを集め、叩打腕識別器を作成する。

## 4.3 叩打支点の識別(要素 3 への対応)

叩打支点を識別するには、図 1 に示すように、腕の関節の動作から判定する。関節座標は Kinect で取得し、関節角度を計算する。関節角度は注目する関節と、それと隣り合う関節の 3 点がなす角であ

る。注目する関節として肩、肘を選ぶ。これにより 3 種類に識別することができる。各叩打の直前フレームにおける関節角度変化量が閾値以上の場合にその関節が動いたとする。この識別を行うには左右どちらの腕の関節情報を参照するかの情報が必要なため、叩打腕の識別ができることが前提となる。

## 5. 提案手法の実装と評価

提案手法の有効性を確かめるためにデータを集め、学習を実行した。まず関節角度だけを入力として、楽器はスネアドラムだけを学習した(実験 1)。次にドラムスティックの位置、演奏情報を加えて、対象楽器を 7 つに増やして学習した(実験 2)。

### 5.1 学習データの構築と検証方法

すべての実験で筆者のひとりが叩打したデータを提案手法の方法で収集した。データは 1 ショットずつ叩打した瞬間の 15 フレーム前までの演奏情報を取得した。学習は機械学習ライブラリの Keras を利用し、モデルは図 2 のように LSTM で構築した。LSTM のユニット数は 50、バッチサイズは 32、エポック数は 1000 とした。訓練用データをとトレーニングデータ：バリデーションデータ=8 : 2 として学習させ、その結果を学習したモデルによる正解率をテスト用データを用いて検証した。

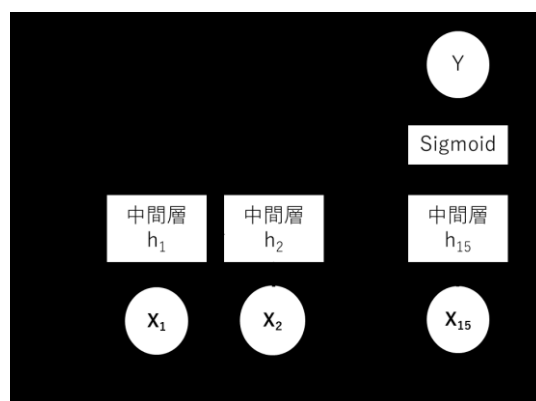


図 2:RNN(LSTM)による叩打腕識別学習モデル

### 5.2 実験 1

図 1①スネアドラムでの叩打を対象として訓練用データを 400 ショット、テスト用データを 200 ショット収集した。叩打腕の左右の割合は訓練、テストどちらも半々である。右手のデータを収集しているとき、左手の姿勢は自由とした。左手でも同様である。入力は腕の関節角度のみである。

このデータを用いて学習を実行したところ、学習履歴は図 3 のようになった。テスト用データを用い

て学習したモデルの識別の正解率を検証した結果、0.72 となった。

### 5.3 実験 2

図 1①～⑦すべての叩打を対象としてそれぞれ訓練用データを 200 ショット、テスト用データを 100 ショット、計 2100 ショット収集した。叩打腕の左右の割合は訓練、テストどちらも半々である。右手のデータを収集しているとき、左手のドラムスティックはなるべく叩打対象から離すようにした。左手も同様である。入力は腕の関節角度、ドラムスティックの位置、叩打面の種類、叩いた強さとした。

このデータを用いて学習を実行したところ、学習履歴は図 4 のようになった。テスト用データを用いて学習したモデルの識別の正解率を検証した結果、0.90 (小数第 3 位で四捨五入) となった。

### 5.4 考察

実験 1 よりも実験 2 で識別精度が向上した。図 3 よりも図 4 でバリデーションの精度が向上している傾向もみられる。どんな要因で向上したか考察する。これはデータ収集の際の体の姿勢が大きな要因であると考えられる。そう考える理由は 3 つある。第 1 の理由として実験 2 では姿勢に制約条件を加えたことを挙げる。これにより姿勢のパターンを学習することができたと考える。第 2 の理由として実験 2 のデータでドラムスティックの情報のみを入力として学習させた場合と、関節角度で入力して学習させた場合では特に識別精度に変化がなかったことを挙げる。前者は約 0.84、後者は約 0.85 となった。第 3 の理由として実験 2 のデータと実験 1 のデータの数や入力次元数を合わせた場合でも実験 2 のデータのほうが識別精度が高かったことを挙げる。データを腕の関節角度、スネアのみにし、訓練を 200 ショット、テスト 100 ショットにしたところ識別精度は実験 1 で 0.86、実験 2 で 0.96 となった。すなわち、データの増加が原因ではないと考える。

この結果は RNN の特性である時系列情報を考慮した学習が行えているとはいえない。しかし、機械学習で叩打腕識別できる可能性を示した。

## 6. おわりに

本稿では、RNN を用いてドラム演奏の叩打腕を識別する手法を提案した。その結果データの収集方法を工夫することで識別できる可能性を示した。

今後の課題として、これまでの成果をもとに学習システムの再設計、演奏認識手法のより詳細な評価実験を行う予定である。

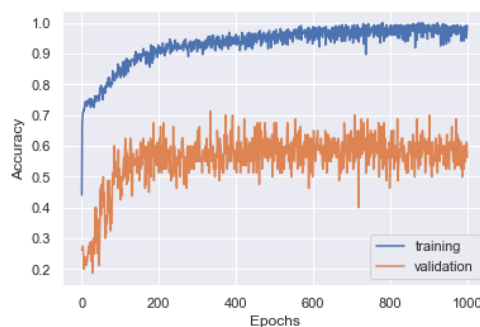


図 3:実験 1 の学習履歴

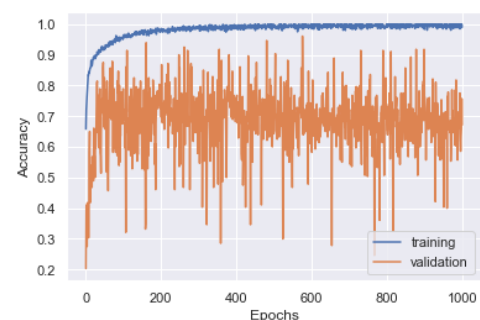


図 4:実験 2 の学習履歴

## 謝辞

本研究はJSPS科研費JP17K01098の助成を受けた。

## 参考文献

- [1] 西村 健次,松原 周信:“ドラム演奏におけるスティックワーク構造の特徴”,京都府立大学学術報告, y「人間環境学・農学」,Vol.59,pp.1-9, (2007)
- [2] 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, Vol.34(1), pp.32-43 (2017)
- [3] Yasuhiro Imada, Youji Ochi,:Hitting Arm Detection for Drum Performance Learning System Using Kinect,Proc. of The Seventh International Conference on Informatics and Applications (2018)
- [4] 今田泰広, 越智洋司, 井口信和: Kinect を利用したドラム奏法の識別手法, 教育システム情報学会全国大会, pp.321-322 (2018)
- [5] 曾我真人,瀧寛和,松田憲幸,高木佐恵子,吉本富士市:スキルの学習支援と学習支援環境(<特集>スキルサイエンス),人工知能学会, Vol. 20.5, pp.533-540 (2005)