

LSTM を用いたジャズの譜面によるアドリブ演奏者判別手法

山住 凌史 指導教員：大久保 弘崇

1 はじめに

現在に至るまで、音楽における個性は感性情報として扱われてきた。具体例として、ジャズの演奏を聴き、演奏の雰囲気や音色から演奏者を判別することが可能でも、その根拠まで詳しく述べることはできない、などが挙げられる。曲を聴き取ることによって曲名や演奏者を判別するアプリケーションは既に存在するが、その手法は既存の波形データから、聞き取った波形データと一致、もしくは類似する物を探して判別する、ものである。この手法は同時に演奏している演奏者の音も利用している為、演奏者個人の個性を利用して判別しているわけではない。またデータが存在しない曲は判別できないという欠点がある。

もし波形データではなく、過去の演奏を記録した楽譜データを用いて、新たな演奏を記録した楽譜データから演奏者を判別することが可能になれば、演奏者の個性が計算機上で扱えるということになる。なぜなら楽譜データというものは、音色や雰囲気などの表現がデータ化されておらず、音階と音が発生した時点、音が持続した時間のみがデータ化されているからである。楽譜データを判別可能ということは、演奏者の音選び、つまり個性のみで演奏者の判別が可能であると言える。音楽における個性を計算機上で扱えるようにすることで、個性を取り入れたアドリブの生成や、音楽理論の発展など様々な有効活用が考えられ、音楽界に大きな進歩をもたらす。

そこで本研究では、二人のジャズプレイヤーによる演奏情報を楽譜から数値化し、機械学習を行うことで、分類が可能かどうかを検証し、可能であればその精度を高めることを目的とする。

2 ジャズの特徴

ジャズの曲の構成は、大きくテーマとアドリブの二つに分割することができる。テーマの部分では、曲の基本となるメロディ(旋律)を演奏する。ポップスやロックは基本メロディのみで構成されている。アドリブの部分では、演奏者が自由に演奏する。

ここでアドリブとは、曲ごとに決められたコード進行上で、演奏者が自由に演奏する事をいう。同じ曲でも演奏者によって全く違うアドリブになるため、アドリブはジャズの中で最も演奏者の個性が出やすい部分である。

3 LSTM

本研究では、楽譜データを数値化する際、時系列データとして変換する。その為機械学習には時系列データを扱うことが可能な LSTM (Long Short Term Memory)[1]を用いる。LSTMとは、同じく時系列データを扱うモデルである RNN (Recurrent Neural Network) を改良したモデルであり、RNN の勾配消失問題を解決すべく考案された。勾配消失とは、時系列データ長に比例してネットワークが深くなることによって勾配が消失し、学習がうまく進まないことである。LSTM では RNN に CEC (Constrant Error Carousel) 及び忘却ゲート (forget gate) を追加し、過去データを保持及び消去することで、長期の時系列データに対応した。

4 手法

最初に二人の演奏者のアドリブを比較する為、同じコード進行上での演奏者ごとのアドリブを収集する。第2に、収集したアドリブのキーを全て揃えることで正規化を行う。第3に、正規化したアドリブを数値化し、学習用データと検証用データに分割する。第4に、学習用データを用いて LSTM で学習する。最後に学習したモデルと検証用データを用いて演奏者の分類を行い、精度を評価する。

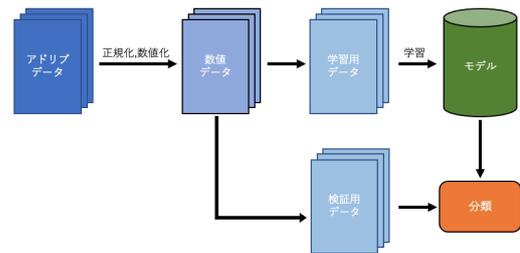


図1: 手法の流れ

5 評価結果と考察

本研究で用いたデータの数は一般的な機械学習に用いられるデータ数よりも少ない為、評価方法として交差検証を用いた。検証結果を図2、図3に示す。

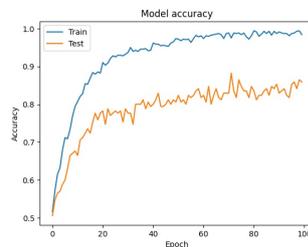


図2: accuracy

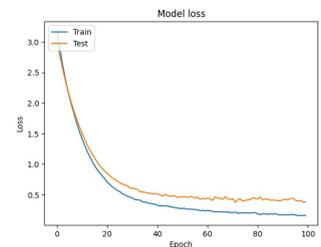


図3: loss

結果から、約8割の未知データに対して正しく分類可能であることがわかった。accuracy のグラフから見て取れる微量の上下変動は、データの個性の度合いによる分類精度の変動だと考えられる。

6 終わりに

本研究では、楽譜データに基づく演奏者分類の可能性を検証した。データによって精度に多少の誤差はあるが、楽譜データから演奏者を分類すること、すなわち個性を分類することは可能であると考えられる。

参考文献

- [1] Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Computation, Vol. 9, No. 8, pp.1735-1780 (1997)