

コード構成モデルの提案と音楽生成システムの設計

2019SE027 小西 遥人

指導教員：野呂 昌満

1 はじめに

近年、機械学習の発展と共に、機械学習を用いた自動作曲への関心が高まっている。自動作曲で利用されるニューラルネットワークは多岐に渡るが、ほとんどの研究でRNN(Recurrent Neural Network) が用いられる [1]。これは RNN が時系列データを扱うタスクに効果的とされていることに起因する。

メロディは単音で構成されるので、知識は必要だがある程度直感的に作ることができる。一方で、コード進行は複数の音の組み合わせで構成され、作成には楽典などの多数の専門知識が必要とされる。以上のことから、メロディに比べて作成難易度が高い。メロディの自動生成が重要な課題であると同様に、コード進行の自動生成も作曲支援の観点から重要な課題である。

現在、機械学習を用いてコード進行を生成、提案するシステムは多数存在するが、それらはダイアトニックコードのみであったり、メジャー・マイナーコードのみを対象としていることが多く、扱えるコードの種類が少ない問題点がある。この問題点には以下の2点の要因がある。

要因 1. 多種多様なコードを扱うためのコードの表現手法が確立されていない。

要因 2. 複雑なコードのコードのサンプル数が少ない。

本研究の目的は、扱えるコードの種類が少ないという問題点を解消し、「多種多様なコードから構成されるコード進行」を生成可能なシステムを設計することである。

この目的を達成するために、問題点の要因 1 への対策として本研究で新たなコードの表現手法であるコード構成モデルを提案する。また、提案モデルに基づく音楽生成システムの設計を行うことによって、目的達成のためのアプローチを行う。その際、設計したシステムが適切にコード進行の生成が可能であるかという観点から妥当性の考察を行う。

研究課題を次の3点とする。

1. コード構成モデルの提案.
2. コード構成モデルに基づく音楽生成システムの設計.
3. 設計したシステムの妥当性評価.

2 関連技術と既存研究

2.1 RNN

RNN とは、再帰的構造を取り入れたニューラルネットワークのことである。再帰的構造により、音楽生成や自然言語処理などの時系列データを扱う際に効果的に作用する。RNN の一種に LSTM(Long Short Term Memory) がある。LSTM は、従来の RNN における勾配消失問題を解消し、長期的記憶を可能にしたものである。

2.2 DeepBach

Gaetan Hadjeres ら [2] は、分類器にニューラルネットワークを用いた疑似ギブスサンプリングを用いて、J.S.Bach のようなポリフォニックな音楽を生成する DeepBach というシステムを開発した。このシステムには、アルゴリズムの制約条件を変更することによって、ユーザーが介入する余地があり、インタラクティブな作曲が可能という特徴がある。

ポリフォニックな音楽の生成モデルであり、提案するモデルとの親和性が高いという理由から、本研究ではこの DeepBach を元に音楽生成システムを設計する。

3 コード構成モデルの提案

本研究では、コードを7つの構成音特徴によって表現するコード構成モデルを提案する。コード構成モデルの概要を表1に示す。

本モデルは全てのコードを root, triad, 7th, 9th, 11th, 13th, bass の7つの特徴量からとらえるというものである。本モデルは、コードネームの命名規則や楽典の和音の項目を参考に、多種多様なコードが表現できるように決定した。

表1 コード構成モデル

特徴名	特徴量												
根音 root	I	I #/ II ♭	II	II #/ III ♭	III	IV	IV #/ V ♭	V	V #/ VI ♭	VI	VI #/ VII ♭	VII	
3和音 triad	major	minor	dim	aug	sus4	sus2							
7度音 7th	null	6th	m7th	M7th									
9度音 9th	null	♭ 9th	9th	#9th									
11度音 11th	null	11th	#11th										
13度音 13th	null	♭ 13th	13th										
ベース音 bass	I	I #/ II ♭	II	II #/ III ♭	III	IV	IV #/ V ♭	V	V #/ VI ♭	VI	VI #/ VII ♭	VII	

4 音楽生成システムの設計

本研究で設計する音楽生成システムは、DeepBach のアーキテクチャを元に設計する。また、音楽生成システムは生成物にユーザーの意図を反映させる機能が重要であるという理由から、ユーザーが入力したコード進行をリハーモナイズするというユースケースを想定する。その際、ユーザは単純な 3 和音のコードを入力するという前提の下で設計を行う。

本研究で設計するシステムのデータの流れについて図 1 に示す。

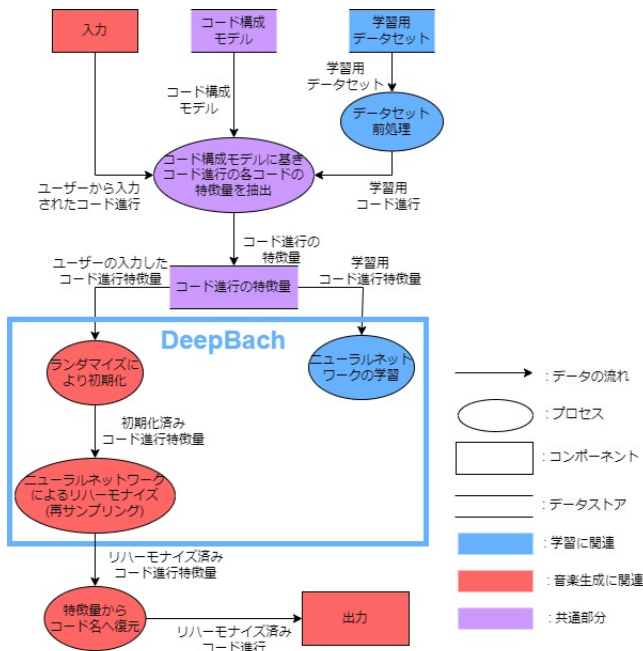


図 1 提案するシステムの処理の概要

本システムでは、ユーザーによって入力されたコード進行から、それぞれのコードについてコード構成モデルに基づいた特徴量を抽出する。その後、root, triad, bass 特徴以外の特徴を乱数によって初期化を行い、特徴量をサンプリングを行うことによってリハーモナイズをする。最終的にリハーモナイズされた特徴量はコードネームに復元され、コード進行として出力を行う。この乱数による初期化と、ニューラルネットワークを用いたサンプリング処理による音楽生成は、DeepBach[2] で提案された手法を用いる。

サンプリングではニューラルネットワークを用いて図 2 で示す推測タスクを行う。推測タスクでは、目的とする特徴量を 1 点選び、過去に関する特徴量・現在に関する特徴量・未来に関する特徴量の 3 つの情報から目的とする特徴量の推測を行う。ランダム化された特徴量に複数回サンプリングを行うことによって、全体を周りの特徴量を考慮した特徴量に更新し、音楽に違和感のない特徴量に収束させる。各特徴につき図 2 のモデルのニューラルネットワークが 1 つ対応するので、本システムには合計 7 つのサンプリング用ニューラルネットワークが存在する。

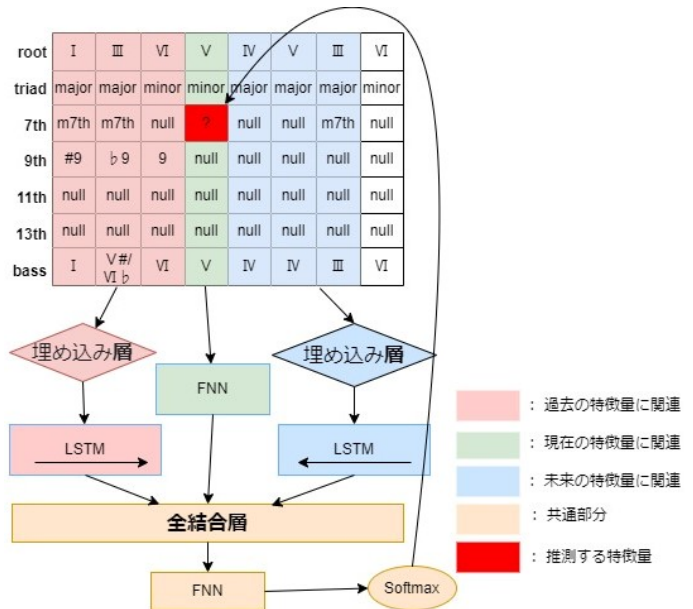


図 2 推測タスクの概要

5 考察

DeepBach[2] では、サンプリングによって 4 つのパートから構成される音楽の生成を行っており、複数のパートから複数のパートの生成が可能である。本研究では提案するコード構成モデルを DeepBach に適用し、パート数を 7 つに拡張し特徴量のサンプリングを行うことによって複雑なコード進行の生成が可能であることを確認した。

本研究ではリハーモナイズを行うことを想定しての設計を行ったが、DeepBach では完全な新規生成も可能である。よって、ソースコードを精査し DeepBach に適切にコード構成モデルを適用することで、リハーモナイズだけでなくコード進行の新規作成や、続きのコード進行を生成するなどの柔軟な使い方もできると考える。

6 おわりに

本研究では、「多種多様なコードから構成されるコード進行」を生成可能なシステムを設計すること目的に、コード構成モデルを提案と音楽生成システムの設計を行い、その妥当性に関して考察をした。

今後の課題として、実際に提案手法に基づいて音楽生成システムの実装を行い、生成されたコード進行についてのアンケート調査を行う必要がある。またその結果をもとにコード構成モデルの妥当性に関してさらなる定性的議論を行い、コード構成モデルが機械学習において効果的に作用するかの評価を行う必要がある。

参考文献

- [1] J.P.Briot,G.Hadjeres,F.D.Pache, “Deep Learning Techniques for Music Generation,” *Springer*, 2020.
- [2] G.Hadjeres,F.O.Pachet,F.Nielsen, “DeepBach: a Steerable Model for Bach Chorales Generation,” *ArXiv*, 2017.