

Web 上の演奏データを利用した作曲支援システム

吉原一期[†]

近山隆[‡]

東京大学 新領域創成科学研究科[†]

東京大学 新領域創成科学研究科[‡]

1. 研究の目的

人間が音楽創作をするときに、全くの無から作品を生み出すわけではなく、過去に聴いた音楽や理論等を利用している。人間が様々な作品を聞き、そこから知識・経験として音楽的素養を蓄え作曲に利用する事をシミュレートすることは、計算機の音楽への関わりを広げる事に繋がると考えている。

本稿では旋律に対する和声付けという行為に焦点を絞り、大量の雑多な音楽データから有意な情報を取り出し、取り出した情報を利用した半自動和声付けシステムを作成する。具体的にはユーザが入力した旋律に対し、インタラクティブに和声進行の候補を提示するシステムを作成する。

大量の雑多なデータを利用するため、データより有用な情報を取り出さなければいけない点、また様々なタイプの楽曲をデータとして使用しなければならない点、この2点が課題となる。

本研究においては、音源データからの採譜技術の発展を見越し、採譜済みの演奏データとして Web 上から収集した MIDI データを元に学習する事を考える。

2. データの収集

Web 上から演奏データを収集するためにクローラーを作成した。MIDI ファイルを公開しているサイト専用の登録式検索サイトなどをルートに、深さ打ち切りで Web を検索していき MIDI ファイルもしくは中に MIDI ファイルがあるアーカイブを収集した。2日ほど動かすことで、2万件弱の MIDI ファイルを取得することができた。収集したファイルは、ハッシュ値を調べることでバイナリ的に同一の物は除外してあるが、同じ曲を演奏したデータもある。

3. 各演奏データの分析

本研究では和声付けを目的とするので、まず、

A music composition support system that uses data of music performances on the Web

[†] Kazuki Yoshihara, School of frontier Sciences, the University of Tokyo

[‡] Takashi Chikayama, School of frontier Sciences, the University of Tokyo

演奏データから和声進行を取り出すことを目的とする。

小節ごとなど単位を決めて、その中で和声の構成音がなっていれば、そのコードに点数をつけて、一番高い点数の和声とその範囲で演奏されていると判定するという手法を応用する。

和声変化のタイミングを判定するため、和声が変化するごとにマイナスの点数をつけ、一拍ごとにどういった和声を振るのが、全体として一番高い点数がつくか、前向きアルゴリズムで、求めることで決定した。

ポイントは和声の構成音について一律に均一に与えるのではなく、演奏データに和声名を付加した学習データを用意し、例えばマイナーコード等の場合、ルート音に対して何度の音が使われやすいかという情報と、各種類の和声の出現頻度について学習した。

ルート音に対する相対音で何度の音が使われているかの割合 p から、和声名の判定において与えるポイントを次の式で決定する。

$$\log_{12} \left(\frac{p + \mathbf{d}}{1 + 12\mathbf{d}} \times \frac{1}{12} \right) \quad \text{ただし } \mathbf{d} = 1/(12 \times 13)$$

また、あまり使われない種類の和声名が出すぎてしまうという問題を回避するため、和声の種類ごとに、学習データ内での出現割合 q から係数を求め、上で求めたポイントに乗ずる。係数は $q^{\mathbf{I}}$ とし、 \mathbf{I} は出現頻度が低い和声かどの程度現れるか値を変えて何度か試行し、結果をみて 0.1 に決定した。

結果を評価するため、和声名を付けた学習データを二つにわけ、一つを学習用、もう一つを評価用とした。そして全体の時間のうち何%が同じ和声と判定しているかを調べた。出現割合を利用することで、若干精度があがる事が確認できた (80% 85%)。

また、和声進行を学習する上で、調性も重要となってくる。MIDI フォーマットでは、メタデータとして調性を与える方法はあるが、機械的に収集されてきたデータに調性が正しく与えられているという保証はない。そこで、調性が正しく与えられているデータから、他データの調性を判断す

る事を考える。

演奏データ作成ソフトのデフォルトとして八長調がセットされる事が多いため、八長調以外のものが指定されていた場合、そのデータにおいて調性が正しく設定されていると考え、これらのデータを全て、八長調になるようにシフトさせ、和声進行のモデルを tri-gram で作成、また転調が起こる頻度を学習した。

HMMにより旋律から和声を推定し、その和声進行から調性・転調を判定する手法[1]を用い、調性を転調も含め推定した。

4. 楽曲のクラスタリング

HMM を用いた和声付け手法[1]を用いて、得られたデータを利用し和声付け実験を行った。演奏データを全て八長調にシフトさせ、和声進行についてはトライグラムで学習を行った。しかし結果、基本的な進行しか選ばれなかった。これは、様々なタイプの楽曲を一緒にして学習したため、共通であられるような規則のみしか学習できなかったためと考えられる。

本稿では、あらかじめ進行に特徴がある楽曲同士をクラスタリングしておくことで、様々なタイプの和声付けを実現する事を目指す。まずクラスタリング手法について述べる。

特徴量ベクトルとしては、各和声名の出現確率・一回の進行(メジャーコードから、2度上のマイナーコードに移るなどといったものを一つのフレーズとして考え、フレーズの出現頻度を求める)、各コードのときルート音に対して何度の音が使われるかという要素を持たせた。

特徴量ベクトルから距離が求まり、クラスタ間距離は重心の距離の平均と定義する。その上で初期状態では各楽曲が一つのクラスタと考え、クラスタ間距離がもっとも近い一組を順に一つのクラスタとしてまとめていくという手法でクラスタリングを行った。しかし、この方法を用いると、一つのクラスタが大きくなりすぎてしまう傾向があったため、クラスタに含まれている要素の数と距離から得点を求め、その得点が高いもの同士をくっ付けていくという方法で、一つのクラスタに要素が集まりすぎないように調整を行った。

5. クラスタリング結果を用いた和声付け

一つの旋律に対して様々な和声をつけることが可能であり、必ず正しいと言えるものはない。本研究ではユーザのイメージにあったものを提供するための和声付けシステムを提案する。

提案システムでは、システム側で短い単位で和声進行の候補を提示し、ユーザ側で候補の各進行

を試聴し、イメージに合うものを選択するという形をとる。

和声進行の候補を提示する際に、その時点までのユーザの選択から、現在和声付けしている曲はどのクラスタに属している確率が高いかを判定する。その上でクラスタのデータを重み付けして使うことで、ユーザが候補の中からイメージが合うものを選択していくうちに、ユーザの意図を取得しイメージに近い和声が上位に提示されるようになることを目標とする。

今まで確定した旋律と和声進行 pre がクラスタ n から生成される確率を $likelihood_n(pre)$ とし、クラスタ n から、次の和声 c が生成される確率を $p_n(c|pre)$ とすると、次に和声 c が選択される確率は次式で求める。

$$\sum_n likelihood_n(pre) \times p_n(c|pre)$$

6. 評価実験と課題

クラスタリングを用いた和声付けをする事で、学習データ中にない曲(データ収集した時点以降に作られた曲)について、曲の最初から順番に和声付けしていく事で評価した。

次の和声候補で正解の候補が何番目に提示されるかを調べた。正解候補が156種類の和声の中で5位以内に提示される確率は、データ全てを均等に利用した場合が32%なのに対し、クラスタリングした結果を用いた場合が47%となった。クラスタリングを用いることで予想の精度があがっていることが確認できる。

また、上でのべた手法でクラスタリングを行うのではなく、全くランダムにクラスタリングした結果を用いて同様の実験を行った。この場合はデータを全て均等に利用した場合よりも予想の精度が落ちた。これは4章で述べたクラスタリング手法による結果は、和声付けをする上で意味のあるものになっている事が確認できる。

また、クラスタリング手法を変える事が最終的な予想精度に大きく影響してきた。現在、クラスタリングのための特徴量ベクトルを修正し、よりよいクラスタリング手法を模索している。また、この手法を用いた半自動和声付けシステムのユーザインターフェースを作成中である。

参考文献

[1] 川上隆, 中井満, 下平博, 嵯峨山茂樹, “隠れマルコフモデルを用いた旋律への自動和声付け”, 平成11年情報処理学会音楽情報科学研究会, 99-MUS-34, Jan. 2000.